



## Original Article

**A multi-objective optimization framework for ill-posed inverse problems**

Maoguo Gong\*, Hao Li, Xiangming Jiang

*Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, No. 2 South TaiBai Road, Xian 710071, China*

Available online 20 October 2016

**Abstract**

Many image inverse problems are ill-posed for no unique solutions. Most of them have incommensurable or mixed-type objectives. In this study, a multi-objective optimization framework is introduced to model such ill-posed inverse problems. The conflicting objectives are designed according to the properties of ill-posedness and certain techniques. Multi-objective evolutionary algorithms have capability to optimize multiple objectives simultaneously and obtain a set of trade-off solutions. For that reason, we use multi-objective evolutionary algorithms to keep the trade-off between these objectives for image ill-posed problems. Two case studies of sparse reconstruction and change detection are implemented. In the case study of sparse reconstruction, the measurement error term and the sparsity term are optimized by multi-objective evolutionary algorithms, which aims at balancing the trade-off between enforcing sparsity and reducing measurement error. In the case study of image change detection, two conflicting objectives are constructed to keep the trade-off between robustness to noise and preserving the image details. Experimental results of the two case studies confirm the multi-objective optimization framework for ill-posed inverse problems in image processing is effective.

Copyright © 2016, Chongqing University of Technology. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** Ill-posed problem; Image processing; Multi-objective optimization; Evolutionary algorithm

**1. Introduction**

Ill-posed problems were proposed by Hadamard in the beginning of last century [1]. A problem is defined to be ill-posed if there is no unique solution. Hadamard first defined a mathematical problem to be well-posed when its solution (i) exists; (ii) is unique and (iii) depends continuously on the initial data; otherwise, a problem is ill-posed. An arbitrary small perturbation of the data can cause an arbitrary large perturbation of the solution. Most of the classical physics problems are well-posed. However, their inverse problems are usually ill-posed. Today there is a vast amount of literature on

ill-posed problems arising in many areas of science and engineering [2–5].

Many image processing problems are ill-posed in the sense of Hadamard, such as denoising, deblurring, inpainting and so on. Most of ill-posed problems are not sufficiently constrained. In recent decades, a lot of generic constraints on the problem were introduced to regularize them and make them well-posed. One of the best way to “cure” ill-posed problems is to transform them to well-posed ones by adding (one or more) regularization terms [6]. The regularization methods construct approximate solutions of ill-posed problems that are stable under small changes in the initial data. The problems of finding an approximate solution in regularization methods are to find an appropriate regularizing operator and to determine the regularization parameter  $\alpha$  from supplementary information pertaining to the problem. Tikhonov regularization is one of the most common and well-known form of regularization

\* Corresponding author.

E-mail address: [gong@ieee.org](mailto:gong@ieee.org) (M. Gong).

Peer review under responsibility of Chongqing University of Technology.

methods [6]. The penalty term in Tikhonov function takes into account the additional information about the solution and its role is to stabilize the problem and to single out a useful and stable solution. Besides Tikhonov regularization, there are many other regularization methods with properties that make them better suited to certain ill-posed problems [7–9]. For example, the total variation (TV) model has been introduced by Rudin-Osher and Fatemi (ROF) in [7] as a regularization method. TV operator has been used extensively and with great success for inverse problems because it is able to smooth noise in flat areas of the image. Furthermore, it is a nontrivial application-dependent task to choose suitable parameter values. A large  $\alpha$  favors a small solution seminorm at the cost of a large residual norm, while a small  $\alpha$  has the opposite effect. The choice of the admissible value of parameter depends on the information available with respect to the approximate initial information. Many quantitative parameter optimization methods have been proposed in recent decades, such as the L-curve method, generalized cross-validation, the discrepancy principle and estimation of mean-squared error.

In many cases, the solution to ill-posed problems can be forced to be unique by narrowing the solution space and adding additional information. Another considerable way is to find multiple trade-off solutions and then choose one or more suitable solutions as required. From this respect, it seems that image ill-posed problems can be modeled as problems with incommensurable or mixed-type multiple objectives. Multi-objective evolutionary algorithms (MOEAs) are able to simultaneously optimize multiple objectives to keep the trade-off between these objectives and generate a set of trade-off solutions in a single run [10,11]. For this reason, MOEAs are suitable to solve this kind of problems. The decision makers can judge relatively and select one or more suitable solutions according to the problem requirements.

In this paper, a multi-objective optimization (MO) framework is introduced to solve ill-posed inverse problems in image processing. Sparse reconstruction and change detection are used as two case studies to show how the MO framework can be successfully used to solve image ill-posed problems. Sparse reconstruction is a typical ill-posed inverse problem. In this paper, the measurement error term and the sparsity term are used as the two conflicting objective functions and optimized by MOEAs simultaneously. Sparse denoising of natural images and sparse unmixing of hyperspectral data are implemented to validate the effectiveness of the proposed MO framework. In the case study of change detection in synthetic aperture radar (SAR) images, two conflicting objectives are designed to preserve details and restrain noise, respectively. Experiments on simulated and real data sets demonstrate the superiority of the proposed MO framework.

The rest of this paper is organized as follows: In the next section, the MO framework for image ill-posed problems is introduced. Section 3 shows the case study of sparse reconstruction in detail. The case study of image change detection is described in Section 4. Finally, concluding remarks are given in Section 5.

## 2. MO framework for ill-posed inverse problems in image processing

### 2.1. Introduction to multi-objective optimization

Without loss of generality, the multi-objective optimization problems (MOPs) can be set for minimization. An MOP with  $m$  decision variables and  $n$  objectives can be described as

$$\begin{aligned} \min \quad & F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_n(\mathbf{x}))^T \\ \text{s.t. } & \mathbf{x} = [x_1, x_2, \dots, x_m] \in \Omega \end{aligned} \quad (1)$$

where  $\mathbf{x}$  is the *decision vector*,  $\Omega$  is the *decision space*,  $F : \Omega \rightarrow R^n$  consists of  $n$  real-valued objective functions and  $R^n$  is called the *objective space*. The *attainable objective set* is defined as the set  $\{F(\mathbf{x}) | \mathbf{x} \in \Omega\}$ . In most instances, the objectives in an MOP are contradictory to each other, which means no point in feasible space can minimize all the objectives simultaneously. Hence, multi-objective optimization [10,12], are designed to find the best trade-off relationship among them simultaneously.

Considering a minimization problem for each objective, it is said that a decision vector  $\mathbf{x}_u \in \Omega$  dominates another vector  $\mathbf{x}_v \in \Omega$  if and only if

$$\begin{aligned} \forall i = 1, 2, \dots, n \quad & f_i(\mathbf{x}_u) \leq f_i(\mathbf{x}_v) \\ \exists j = 1, 2, \dots, n \quad & f_j(\mathbf{x}_u) < f_j(\mathbf{x}_v). \end{aligned} \quad (2)$$

And a point  $x^*$  in  $\Omega$  is called a Pareto optimal solution to Eq. (1) on condition that there is no such point  $x$  in  $\Omega$  that makes  $F(x)$  dominate  $F(x^*)$ . Then  $F(x^*)$  is termed as Pareto optimal vector. The objectives in a Pareto optimal vector have such relationship: a decrease in one objective causes an increase in the others. All the Pareto optimal points constitute a set called Pareto optimal set [13], and their corresponding Pareto optimal objective vectors are called the Pareto optimal front (PF) [13].

For multi-objective optimization, it has been recognized that evolutionary algorithms (EAs) are well suited because EAs can deal with a set of possible solutions simultaneously [11,12]. Since [14], various EAs to deal with MOPs have been proposed and these EAs are termed as multi-objective evolutionary algorithms (MOEAs). MOEAs seek to obtain a set of Pareto optimal solutions for approximating the true PF in a single run. Most of current MOEAs can be classified into three categories. The first category is Pareto dominance based, such as the nondominated sorting genetic algorithm II (NSGA-II) [15] and the simulated annealing-based multi-objective optimization algorithm (AMOSA) [16]. The second one is performance indicator based, the hypervolume is the well-known indicator used in these algorithms. The third category is decomposition based [17,18], these algorithms decompose an MOP into a number of single objective subproblems and optimize them simultaneously. Each subproblem is optimized by only using information of its several neighboring subproblems.

## 2.2. Multi-objective optimization to ill-posed inverse problems

Many image inverse problems are ill-posed, which can be generally modeled as

$$\mathbf{y} = \mathbf{Ax} + \mathbf{v} \quad (3)$$

where  $\mathbf{x}$  is the unknown image to be estimated,  $\mathbf{A}$  is degrading operator and  $\mathbf{v}$  is additive noise. When  $\mathbf{A}$  is identity, the problem becomes image denoising; when  $\mathbf{A}$  is a blurring operator, the problem is deblurring; when  $\mathbf{A}$  is a set of random projections, the problem becomes compressed sensing; when  $\mathbf{A} = \mathbf{D} \cdot \mathbf{H}$  where  $\mathbf{D}$  is a down-sampling operator and  $\mathbf{H}$  is a blurring operator, the problem becomes single image super-resolution.

Due to the ill-posed nature of image inverse problems, the solution to Eq. (3) with an  $l_2$ -norm fidelity constraint is generally not unique, which is described as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{Ax}\|_2^2. \quad (4)$$

In order to find a solution to Eq. (3), prior knowledge of images is used to regularize the ill-posed inverse problems. However, the obtained solution can only be an approximate one at best. While prior knowledge is introduced, it also introduces the regularization parameter. The parameter is an important quantity which controls the properties of the regularized solution. But the regularization parameters are sensitive to different data sets, which is a hard task to choose suitable regularization parameter values.

In this paper, the image ill-posed problem is modeled as an MOP, which is optimized by MOEAs to get a set of trade-off solutions. Because the solution to the ill-posed problem is not unique, we can obtain a group of solutions in a single run and then the decision maker can choose one or more suitable solutions by taking various factors into account instead of tuning regularization parameters to obtain various solutions. However, there still exists several problems in the process of using MOEAs to solve image ill-posed problems.

1) *Whether the image ill-posed problem can be modeled as an MOP?* The solution to many image ill-posed problems is not unique because they have incommensurable or mixed-type objectives. It is reasonable to model the image ill-posed problem as an MOP. The selected objectives should be as contradictory as possible. In this paper, this MOP is optimized by multi-objective evolutionary algorithms.

2) *How to overcome the difficulties of dimensionality curse?* In MOEAs, one of the most difficult tasks is to choose appropriate decision vectors. In the field of image processing, this issue becomes more important. Because the number of pixels is large, e.g.,  $100 \times 100$ , if the gray values of all pixels are searched directly using the evolutionary algorithms, the search space dimension of the decision vectors (more than ten thousand) will be very large and the computation will be excessive. Obviously, it is worth encouraging to use advanced techniques to avoid encoding all pixels.

It is an urgent task to select suitable objectives based on the properties of ill-posedness in the beginning of the MO framework for image ill-posed problems. Without loss of generality, bi-objective optimization is considered in this paper. Fortunately, it seems that sparse representation is an effective way to transform the image ill-posed problem into an MOP by providing sparse priors. We can obtain a set of trade-off solutions by simultaneously optimizing the measurement error term and the sparsity term. In order to reduce dimensionality, the sparse coefficients are chosen as the decision vectors, whose dimensionality depends on the size of dictionary. Sparse representation is widely used to solve plenty of image ill-posed inverse problems, such as denoising, deblurring, super-resolution and so on [3–5,19]. This paper chooses the spare denoising problem as the fundamental one. Furthermore, a challenging problem with high-dimensional data, spare unmixing of hyperspectral data, is researched. Section 3 gives a detailed description of the MO framework for sparse reconstruction problems. Besides sparse representation, other priors can also be used for designing the conflicting objectives. In Section 4, two problem-specific objective functions are proposed for change detection in SAR images, which focus on addressing the trade-off between robustness to noise and effectiveness of preserving the details. In order to avoid difficulties of dimensionality curse, fuzzy clustering technique is used by encoding the cluster centers instead of all pixels. We decompose this MOP into a number of scalar optimization subproblems with different weight values and optimize them with evolutionary algorithms. A number of solutions representing different trade-off relationships between preserving details and restraining noise are given by the proposed method. The decision makers can judge relatively and choose one or more appropriate solutions based on the available information.

## 3. Case study: sparse reconstruction

### 3.1. Ill-posed sparse reconstruction

Sparse reconstruction algorithms have been studied for nearly a century, and they arise in many areas and have found many applications in image processing. Considering an under-determined system of linear equation  $\mathbf{y} = \mathbf{Ax} + \mathbf{n}$ , sparse solutions are obtained by optimizing the following equation:

$$\begin{aligned} \min_{\mathbf{x}} & \|\mathbf{x}\|_0 \\ \text{s.t. } & \|\mathbf{y} - \mathbf{Ax}\|_2^2 \leq \sigma \end{aligned} \quad (5)$$

where matrix  $\mathbf{A} \in R^{M \times N}$  with  $M \leq N$ ,  $\mathbf{y} \in R^M$  and  $\mathbf{x} \in R^N$ .

It has been proven to be NP-hard optimization problems to solve the above equation. Researchers have proposed many approaches to deal with this problem, such as orthogonal matching pursuit [20], iterative hard thresholding methods [21], relaxation algorithms [22] and so on. Among them, a widely used method is considering the following the  $l_1$ -minimization problem:

$$\min_{\mathbf{x}} \frac{1}{2} \left\| \mathbf{A}\mathbf{x} - \mathbf{y} \right\|_2^2 + \lambda \left\| \mathbf{x} \right\|_1 \quad (6)$$

These works relax the non-convex  $l_0$ -norm with the convex  $l_1$ -norm to construct approximate solutions. Obviously, it is a hard task to keep trade-off between reconstruction error and a sparsity requirement. The more sparse the vector  $\mathbf{x}$  is, the higher the measurement error is. Therefore image sparse reconstruction is referred as an ill-posed inverse problem and can be well solved by the MO framework to balance the trade-off between the measurement error and the sparsity terms.

### 3.2. Sparse reconstruction using MO framework

As shown in Fig. 1, the MO framework is utilized to solve the sparse reconstruction problems. As described above, the sparse constraint and measurement error are selected as the two conflicting objectives. Then the sparse reconstruction problem can be modeled as the following MOP:

$$\min_{\mathbf{x}} \mathbf{F}(\mathbf{x}) = \min_{\mathbf{x}} (\|\mathbf{x}\|_0, \|\mathbf{Ax} - \mathbf{y}\|_2^2) \quad (7)$$

In this section, an MOEA with cooperative coevolutionary strategy is used to solve the problem above. The MOEA we used is the nondominated sorting genetic algorithm (NSGA-II) proposed by Deb et al. in [15], which builds a population of competing individuals and ranks them on the basis of non-dominance. Although the coefficient  $\mathbf{x}$  is encoded in sparse representation instead encoding all pixels, the dimensionality

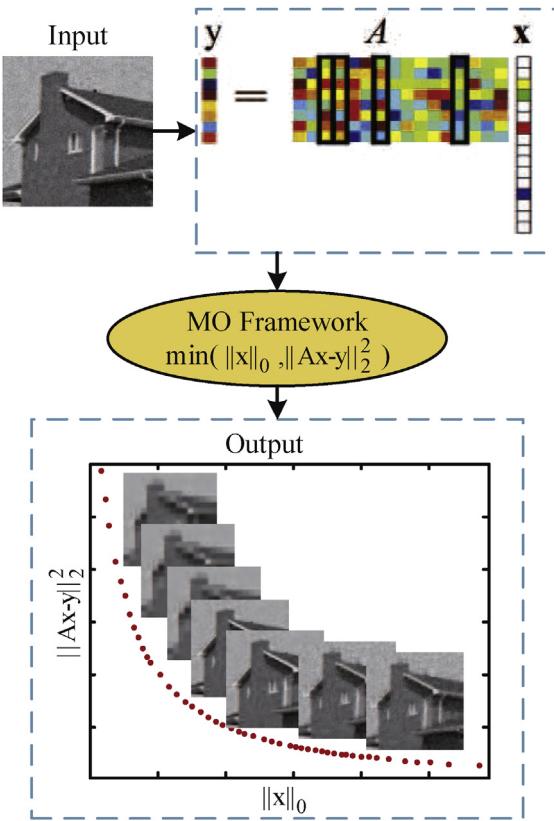


Fig. 1. Sparse reconstruction using MO framework.

is still high. In order to deal with the high-dimensional problems, the cooperative coevolutionary strategy is used in the proposed method. It can be regarded as an automatic approach to implement the divide-and-conquer strategy and has been proved effective to large scale global optimization problems [23–25]. The procedure of the multi-objective evolutionary algorithm for sparse reconstruction is shown in Algorithm 1 [15,26,27,28].

---

**Algorithm 1** A multi-objective evolutionary algorithm for sparse reconstruction

---

- 1: **Step 1)** Set  $t = 0$  and max generation number  $maxgen$ .
  - 2: **Step 2) Initialization:** Generate the initial population  $P_0$  with size  $N_p$ .
  - 3: **Step 3)** Obtain the optimal PF of  $P_0$  and select a solution (named  $P_{sel}$ ) from PF randomly.
  - 4: **Step 4)** Divide each decision vector ( $m$ -dimensional) into  $n$  subcomponents ( $s$ -dimensional) randomly, i.e.  $m = n \times s$ . And then the population is split into  $n$  subpopulation ( $SubP$ ) with size of  $N_p \times s$ .
  - 5: **Step 5)** For each subpopulation  $SubP$ , set  $i = 1$  and the max generation number of collaboration  $maxc$ :
  - 6: **Step 5.1)** Generate a new subpopulation  $SubP_{new}$  with mutation [26] and crossover operators [27] and then combine  $SubP$  and  $SubP_{new}$ , i.e.  $SubP_{com} = SubP \cup SubP_{new}$ .
  - 7: **Step 5.2)** Execute the collaboration operator: Generate the population  $P_{com}$  by forming the collaboration with  $SubP_{com}$  and  $P_{sel}$ .
  - 8: **Step 5.3)** Obtain a new population  $P_{new}$  using elite-preserving operator [15].
  - 9: **Step 5.4)** Extract the subpopulation from  $P_{new}$  to replace the subpopulation  $SubP$ .
  - 10: **Step 5.5)** Termination criteria: If  $i < maxc$ ,  $i++$  and then go to **Step 5.1**. Otherwise, go to **Step 6**.
  - 11: **Step 6)** Implement the local search operator proposed in [28] and then obtain a new population  $P_t$  for the next generation.
  - 12: **Step 7) Stopping criteria:** If  $t < maxgen$ ,  $t++$  and then go to **step 3**. Otherwise, stop the algorithm.
- 

The initial population  $P_0$  is generated randomly or partly by problem-specific method.  $t$  is the current generation number and  $P_t$  is the current population with size  $N_p$ . Each individual  $\mathbf{x} = (x_1, x_2, \dots, x_m)$  in  $P_t$  is an  $m$  dimension vector of decision variables. After initializing the population  $P_0$ , each vector of decision variables of dimension  $m \in N$  is divided into  $n$  sub-components randomly. Each subcomponent is generated from a randomly grouping of decision variables in order to better capture the variable interdependencies. Meanwhile  $n$  subpopulations are created and each one has  $N_p$  individuals. For each subpopulation, the mutation and crossover operators are applied to the current subpopulation to obtain new solutions and then the collaboration operator will be executed. The elite-preserving operator is implemented to obtain the new population  $P_{new}$ . And we will extract the subpopulation form  $P_{new}$

to replace the original subpopulation for the next generation of collaboration. The population  $P_t$  is evolved with the evolution of each subpopulation. In order to further search exhaustively, the local search proposed in [28] are employed to obtain the next generation. In the end of MO framework, optimal trade-off solutions to the sparse reconstruction problems can be obtained by angle-based method [29].

### 3.3. Experiment 1: sparse denoising of natural image

The image denoising problem is one of the inverse problems and provides a convenient platform over which image processing ideas and techniques can be assessed. This section demonstrates the efficiency of the MO framework to a fundamental image ill-posed problem. Fig. 2(a) shows a ground-truth image of a white circle against a black background and (c) represents a set of Haar wavelets of (a). Our proposed method is tasked with reconstructing the noise-corrupted image shown in (b) with corresponding Haar wavelets shown in (d). In this experiment, the original image was corrupted with additive Gaussian white noise to achieve signal-to-noise ratio (SNR) 20:1.

Fig. 3 shows various reconstruction error and measurement error for varying amounts of sparsity. The point **b** is a knee point in the knee areas, where further improvement in one objective causes a rapid degradation in other objectives [29]. The solutions to the left of the knee (such as the point **a**) are over-sparse, and these solutions contain large measurement and reconstruction error. The reconstruction images are of serious distortion because these images contain too much blank-gray points as shown in the upper left of Fig. 3. Conversely, the solutions to the right of the knee (such as the point **c**) are not sparse, and these solutions have small measurement error. The reconstruction images are of light distortion and contain a large number of noise as shown in the bottom right of Fig. 3. As shown in Fig. 3, the measurement error of the reconstruction images monotonously decreases

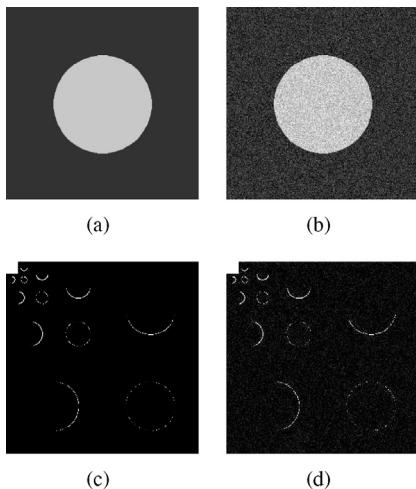


Fig. 2. A synthetic image for sparse denoising: (a) Original image. (b) Image corrupted by noise. (c) Wavelet coefficients of the original image. (d) Wavelet coefficients of the corrupted image.

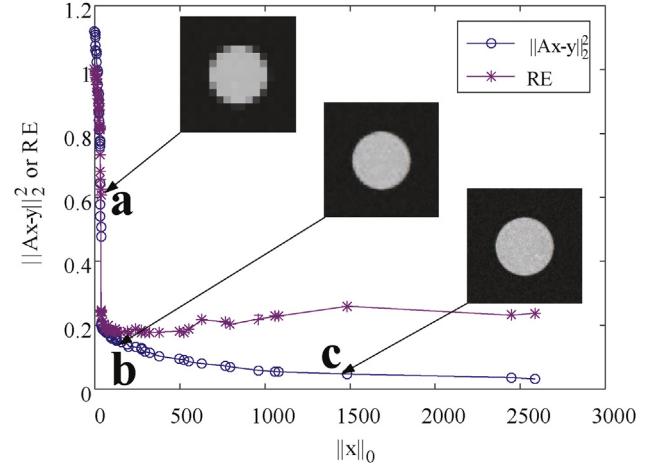


Fig. 3. Variation of reconstruction error (represented by red line) and measurement error (represented by blue line) for varying amounts of sparsity.

with the increase of  $\|x\|_0$ . The reconstruction error of the reconstruction images monotonously decreases firstly and then slightly increases with the increase of  $\|x\|_0$ . Therefore, the two objectives, minimal  $\|Ax - y\|_2^2$  and minimal  $\|x\|_0$ , are contradictory. Obviously, the solutions in the knee region (such as the point **b**) represents the better trade-off between the measurement error and the sparsity, which has the least reconstruction error. The reconstruction image in the knee region is closer to the real image than others. Wavelet coefficients of the solutions in Fig. 3 are shown in Fig. 4.

In order to further assess the MO framework for sparse denoising, the experiments on three natural images are shown below. As shown in Fig. 5, the first column shows the original images and the images corrupted by noise with SNR = 20 are shown in the second column. In the column of L-knee, the reconstruction images are blurred enough because the vector  $x$  is too sparse. Conversely, in the last column of R-knee, the reconstruction images have many noise for the vector  $x$  is not sparse, although they have less measurement errors. By contrast, the solutions in the knee region present better compromise between reconstruction error and measurement error. Fig. 6 shows the reconstruction errors obtained by five sparse reconstruction algorithms, such as homotopy method [30], SpaRSA [31], iterative shrinkage thresholding (IST) [32] and alternating direction method (ADM) [33]. Obviously, the reconstruction errors obtained by the proposed method are less than those obtained by other four algorithms. The MO

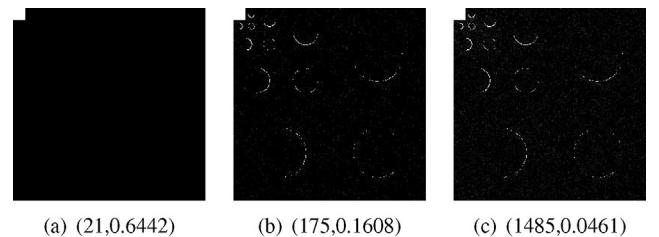


Fig. 4. Wavelet coefficients of the solutions: (a) The left solution of the knee. (b) The knee solution. (c) The right solution of the knee.

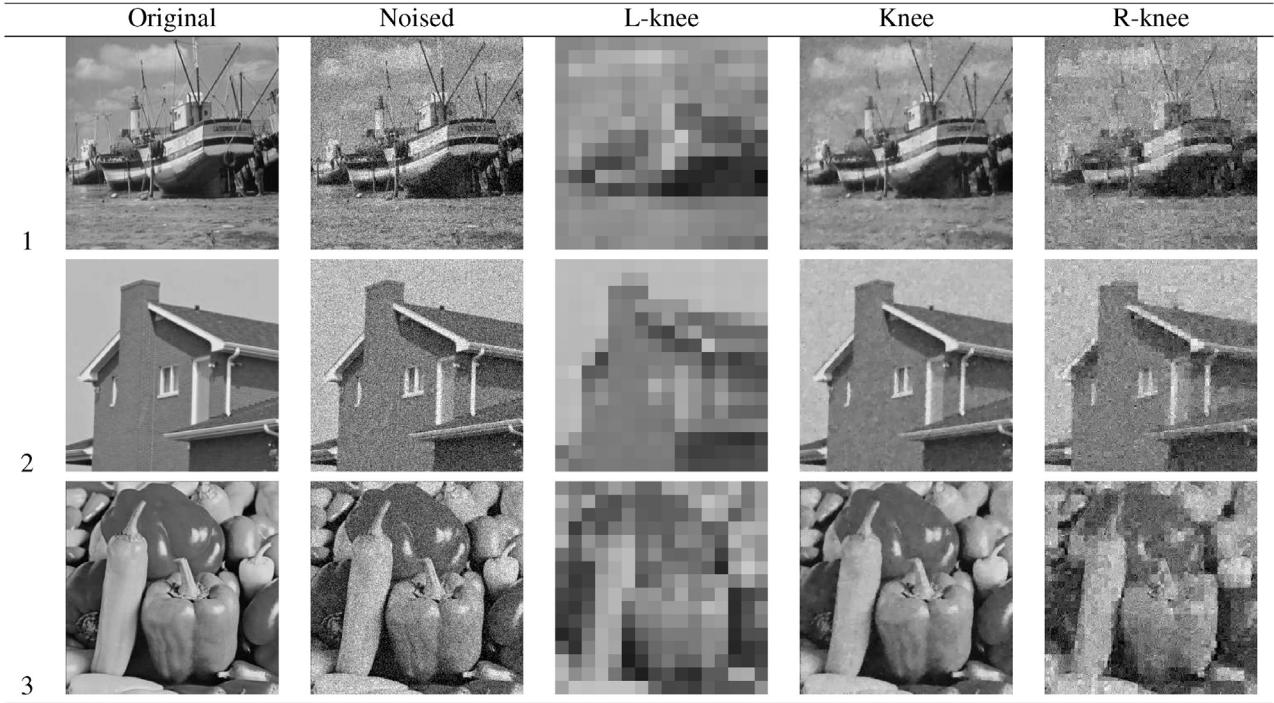


Fig. 5. Sparse denoising using MO framework for three natural images (1: boat, 2: house and 3: peppers). L-knee: the solution has little noise but is overly sparse (the solution on the left); Knee: the solution has the optimal trade-off (the solution in the knee region); R-knee: the solution is not very sparse, which suffers from noise (the solution on the right).

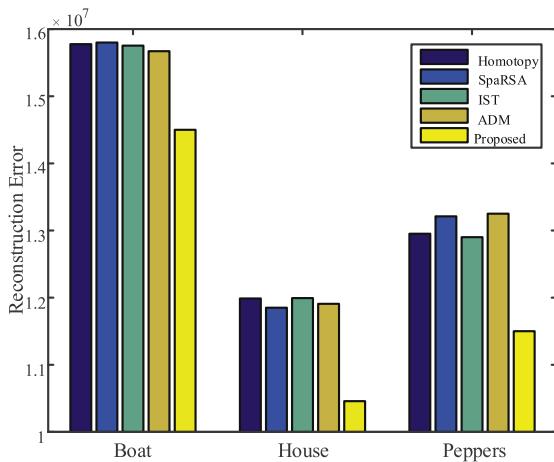


Fig. 6. Average reconstruction errors of five algorithms on three natural images.

framework for sparse denoising gets a better trade-off between the measurement error and the sparsity.

### 3.4. Experiment 2: sparse unmixing of hyperspectral data

Hyperspectral remote sensing is a powerful technique to improve the ability to recognize different landcover classes through a set of images that are obtained over hundreds of contiguous spectral bands. The wealth of spectral information available from hyperspectral imaging instruments has attracted widespread interests because of many applications in various

disciplines, such as mineral exploration, environmental monitoring and military surveillance. The problem of hyperspectral unmixing is one of the most important problems that restrict the development of remote sensing technology, which can be solved by the sparse unmixing technique. Sparse unmixing aims at separating the mixed pixels into a collection of constituent spectra or endmembers, and estimating their fractional abundances. The sparse unmixing model has been widely used in recent years [34–37]. The sparse unmixing via variable splitting and augmented Lagrangian (SUuSAL) [34] based on the alternating direction method of multipliers has been proposed to cope with the sparse unmixing problem. The accuracy assessment used to evaluate the abundance estimations for all the experiments is the signal-to-reconstruction error (SRE) [38], and it is defined as follows

$$SRE(dB) \equiv 10 \log_{10} \left( \frac{E[\|\mathbf{x}\|_2^2]}{E[\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2]} \right) \quad (8)$$

where  $\hat{\mathbf{x}}$  stands for the estimated fractional abundance vector.

The simulated data shown in Fig. 7 contains  $100 \times 100$  pixels and is provided in [38], which exhibits a good spatial homogeneity and has piecewise smooth fractional abundances. Both ANC and ASC are imposed in each simulated pixel in the data set. The data set is also corrupted by different levels of correlated noise (SNR = 10, 20, 30, 40, 50, 60 dB). The unmixing results on hyperspectral data with SNR = 30 dB are shown in Fig. 8. From Fig. 8, we can see that the proposed method obtains a better visual effect and contains fewer noise points. The maps obtained by the proposed method contain

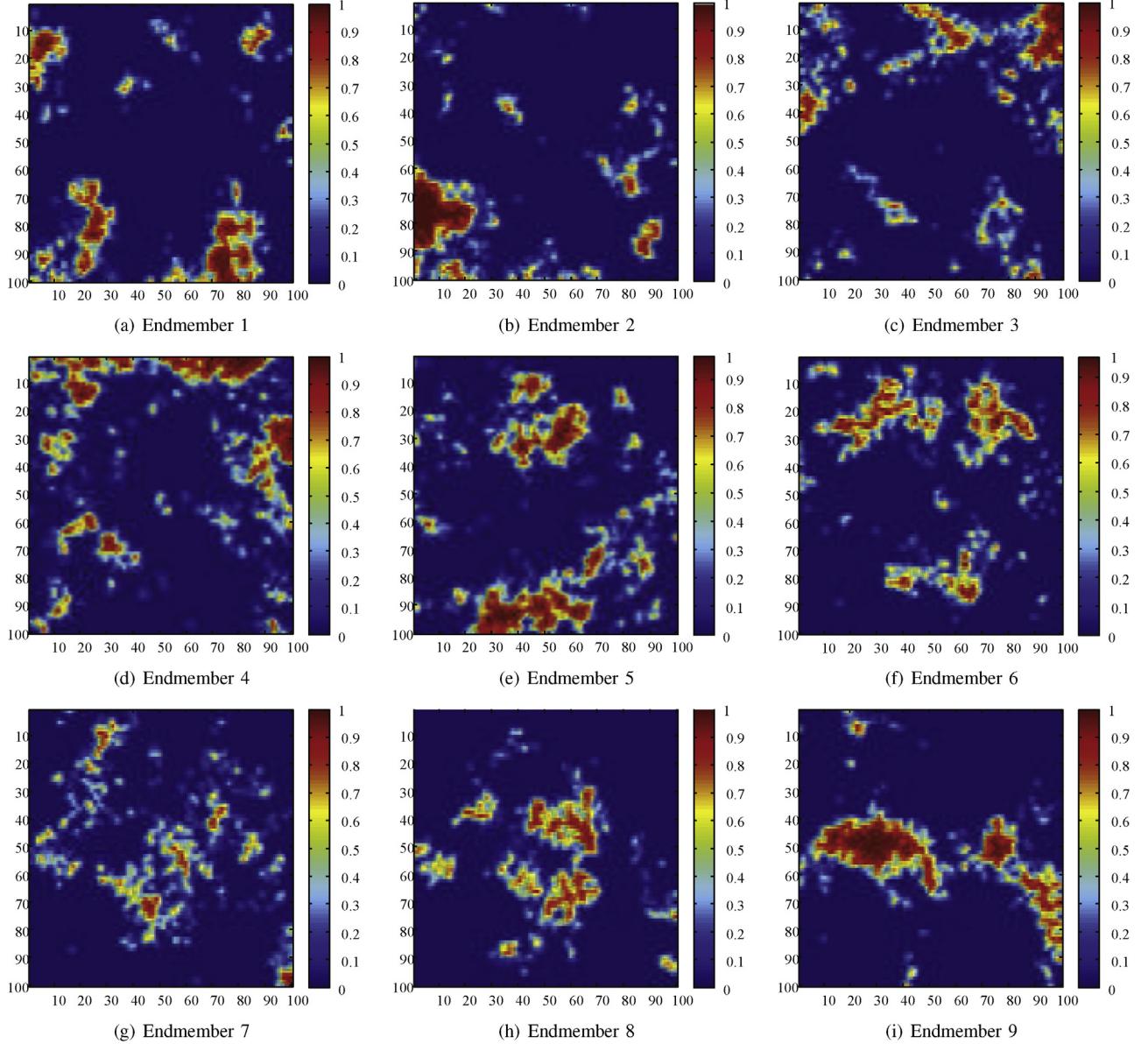


Fig. 7. True fractional abundances of endmembers.

less noise than that of SUoSAL in the rufous regions. Fig. 9 shows the SRE (dB) obtained by different tested algorithms on hyperspectral data set corrupted by different levels of white noise. The values of SRE obtained by the proposed method are a little larger than those of SUoSAL, which validates the effectiveness of the proposed method.

#### 4. Case study: change detection

##### 4.1. Ill-posed change detection

Image change detection analyses two registered images acquired over the same geographical area at different times to identify changes that may have occurred in the study area between the two times considered. Change detection in synthetic aperture radar (SAR) images is full of challenges

because of the presence of the speckle noise in SAR images. For a SAR image, the image intensity  $I$  is related to the underlying backscattering coefficient  $X$  by the multiplicative model  $I = FX$  where  $F$  is the normalized fading speckle-noise random variable. With the log-ratio operator, the multiplicative speckle noise can be transformed in an additive noise component and the range of variation of the ratio image will be compressed and thereby enhances the low-intensity pixels. In [39,40], a multilevel representation of the multitemporal information was computed by decomposing the log-ratio image into several images of the same size as the original one to reduce the noise impact. Bzai et al. [41] used the generalized Gaussian distribution to model the changed and unchanged classes to reduce the corruption of the speckle noise. Furthermore, many probability density functions such as Log normal, Generalized Gaussian, Nakagami ratio, and Weibull ratio were

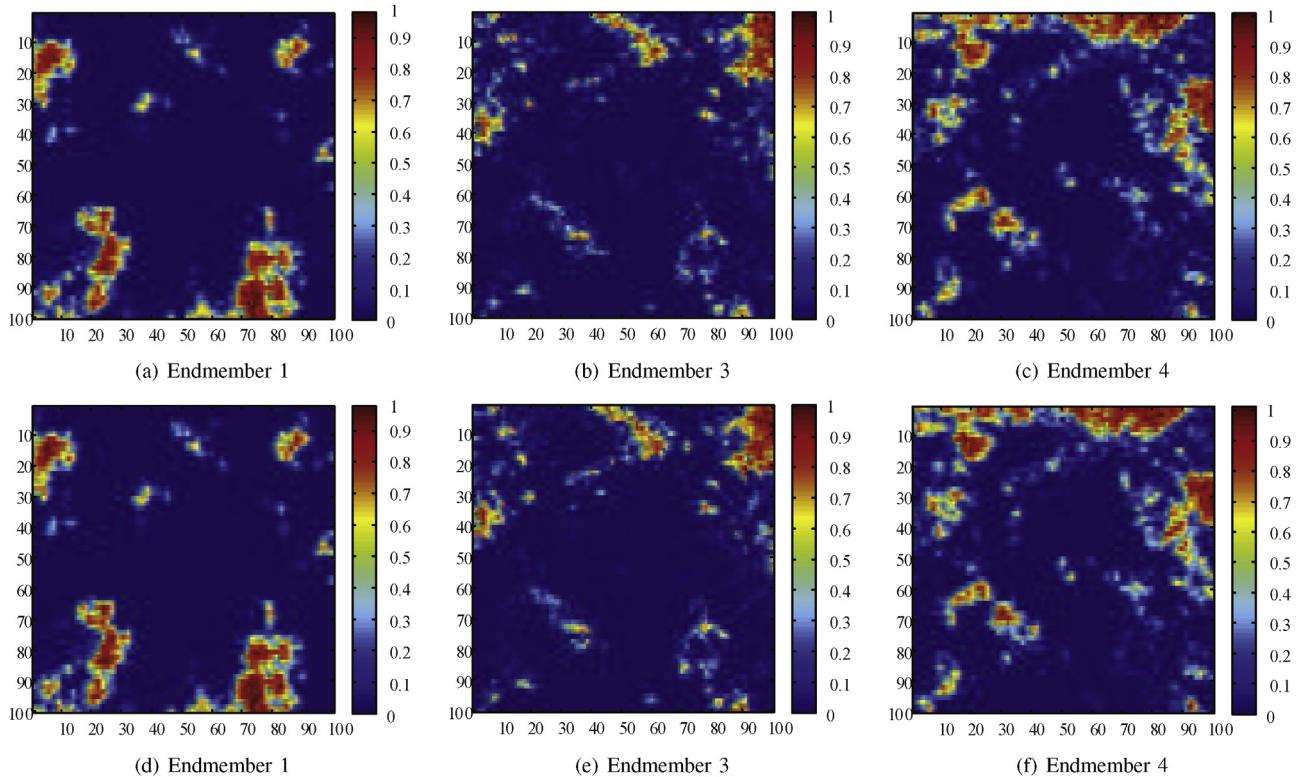


Fig. 8. Abundance maps obtained by different unmixing methods of endmember 1, 3, and 4 (a–c) are obtained by the SUoSAL algorithm (d–f) are obtained by the proposed method.

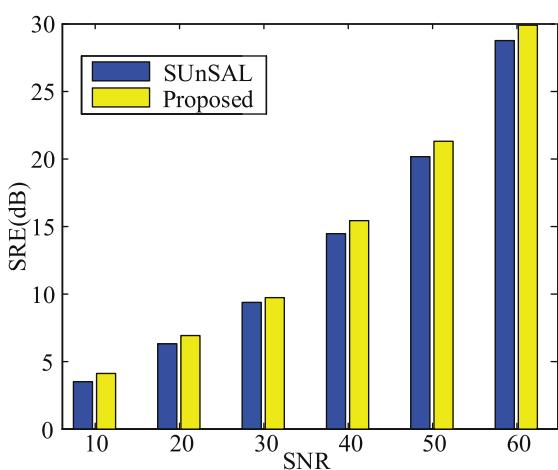


Fig. 9. Average SRE of two sparse unmixing algorithms on hyperspectral data.

investigated to model the distribution of the two classes in [42]. In [43], Hu et al. presented an automatic and effective approach to the thresholding of the log-ratio change indicator whose histogram may have one mode or more than one mode. Moreover, many researchers have added local or non-local information into their algorithms to make them robust to noise. In [44], Gong et al. proposed a reformulated fuzzy local-information C-means (RFLICM) for classifying changed and unchanged regions, which incorporates the information about spatial context by adding a new fuzzy factor into its objective function for the purpose of enhancing the changed information

and reducing the effect of speckle noise. In [45] [46], Markov random field (MRF) based change detection algorithms were proposed to combat speckle noise, which provides a basis for modeling information about the mutual influences among image pixels. Furthermore, non-local means (NLM) method that utilizes nonlocal relationship between image pixels was also used in many change detection techniques to reduce the corruption of the speckle noise [45,47].

However, in most cases, an optimal solution for the image change detection problem does not exist. As shown in Fig. 10(a) is a synthetic SAR ratio image and (b) is the reference image. Two different change detection results are presented in (c) and (d). Some changed regions are interfered by speckle noise, which can not be identified clearly. From the point of restraining noise, the result of (c) seems better than that of (d). But (c) loses some image details because many changed areas are not detected. In the respect of preserving details, it seems that the result of (d) is better although it has a bit of noise. Therefore it is difficult to keep the trade-off between preserving details and removing noise for image change detection problem. In order to solve the problems above, the MO framework is used to tackle the ill-posedness of image change detection. The schematic diagram of the proposed change detection technique is shown in Fig. 11. The log-ratio and the filtered log-ratio images are used to construct the two conflicting objective functions to preserve details and restrain noise, respectively. A set of trade-off solutions are obtained by the proposed change detection method using MO framework.

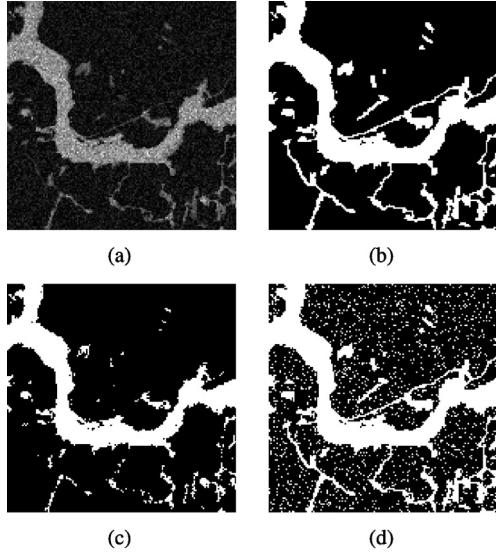


Fig. 10. Schematic diagrams of change detection results. (a) A synthetic SAR ratio image. (b) The reference image. (c) Change detection result with lost details. (d) Change detection result with residual noise.

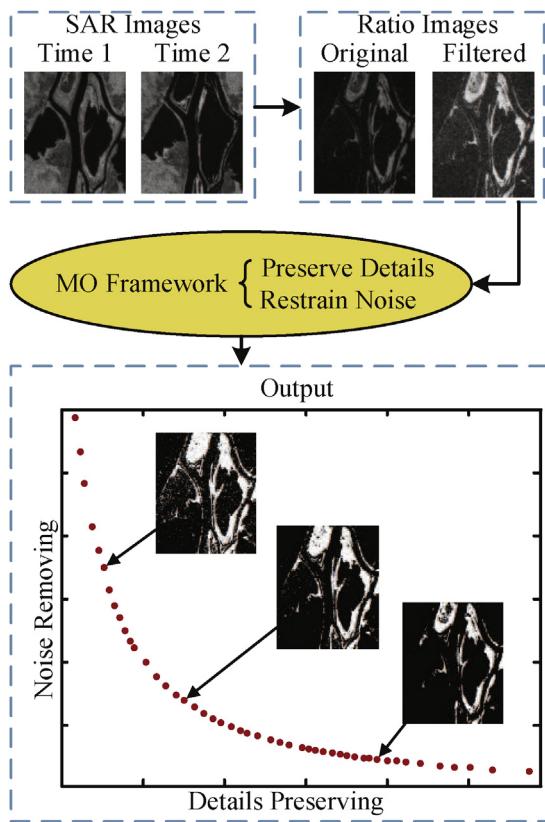


Fig. 11. Change detection using MO framework.

#### 4.2. Change detection using MO framework

As described in Section 2, the choice of objective functions should be as contradictory as possible. In this paper, two objectives based on fuzzy c-means (FCM) measurement are

constructed to preserve image details and restrain noise. The two objective functions  $f_1$  and  $f_2$  can be described as follows:

$$\min F(\mathbf{z}) = \begin{cases} f_1 = \sum_{i=1}^N \sum_{k=1}^2 [u_{ki}^m \|x_i - z_k\|^2] \\ f_2 = \sum_{i=1}^N \sum_{k=1}^2 [u_{ki}^m \|\bar{x}_i - z_k\|^2] \end{cases} \quad (9)$$

where  $x_i$  is the  $i$ th pixel in the original difference image  $I$  and  $\bar{x}_i$  is the  $i$ th pixel in the filtered difference image  $\bar{I}$ .  $\mathbf{z} = (z_1, z_2)^T$  is the decision vector consisting of two cluster centers. And the difference image generated by log-ratio operator is commonly given as follows

$$I = \left| \log \frac{I_B}{I_A} \right| = |\log I_B - \log I_A| \quad (10)$$

where  $I_A$  and  $I_B$  are two coregistered SAR images. And the filtered difference image  $\bar{I}$  is generated by mean filter to smooth the noise in the difference image.

The cluster centers  $z_1$  and  $z_2$  are chosen as the decision vector. MOEA/D [17,18], is utilized to optimize the two conflicting objective functions. Weighted sum approach is used in the proposed technique, which is described as

$$g^{ws}(z_1, z_2 | \lambda) = \lambda f_1(z_1, z_2) + (1 - \lambda) f_2(z_1, z_2) \quad (11)$$

where  $\lambda = (\lambda, (1 - \lambda))^T$  is the weight vector with  $0 \leq \lambda \leq 1$ . The proposed SAR images change detection method is shown in Algorithm 2 (Eq. (16)) [18].

---

#### Algorithm 2 Change Detection based on MO Framework

---

**Input:**  $maxgen$ : max generation number,  $N$ : the number of the subproblems,  $pop$ : the swarm size,  $T$ : neighborhood size,  $EP$ : an external population.

**Output:** A group of binary maps.

- 1: **Step 1) Initialization**
  - 2: **Step 1.1)** Generate a well-distributed weight vectors and then work out the  $T$  closet weight vectors to each weight vector.
  - 3: **Step 1.2)** Initialize the population  $P = \{z_1, z_2, \dots, z_{pop}\}^T$ .
  - 4: **Step 2)** set  $t = 0$ .
  - 5: **Step 3) Cycling**
  - 6: **for**  $i = 1, 2, \dots, N$ , **do**
  - 7:   **Step 3.1)** Obtain the fuzzy membership matrix  $u_i^{t+1}$  by Eq. (16) with cluster centers  $z_i^{t+1}$ .
  - 8:   **Step 3.2)** Generate a new solution with differential evolution and polynomial mutation proposed in [18].
  - 9:   **Step 3.3)** Update neighboring solutions.
  - 10:   **Step 3.4)** Update  $EP$ .
  - 11: **end for**
  - 12: **Step 4) Stopping criteria:** If  $t < maxgen$ ,  $t++$  and then go to **step 3**. Otherwise, stop the algorithm and generate a group of the change detection maps.
-

For image change detection problem, it has two cluster centers for unchanged class and changed class. In this paper, the idea of decomposition is also used in the process of updating the membership values. Decomposition transforms the MOP into a set of distinct scalar aggregation problems. The function of each subproblem is defined as:

$$F_\lambda = \lambda \sum_{i=1}^N \sum_{k=1}^2 [u_{ki}^m \|x_i - z_k\|^2] + (1 - \lambda) \sum_{i=1}^N \sum_{k=1}^2 [u_{ki}^m \|\bar{x}_i - z_k\|^2] \quad (12)$$

with constraints

$$\sum_{k=1}^2 u_{ki} = 1. \quad (13)$$

The fuzzy membership functions are obtained through an iterative process by minimizing Eq. (12) under the constraints defined in Eq. (13). Thus, using the Lagrange multipliers, the energy function is modified as:

$$F_\lambda = \lambda \sum_{i=1}^N \sum_{k=1}^2 [u_{ki}^m \|x_i - z_k\|^2] + (1 - \lambda) \sum_{i=1}^N \sum_{k=1}^2 [u_{ki}^m \|\bar{x}_i - z_k\|^2] + \alpha_i \left( 1 - \sum_{i=1}^2 u_{ki} \right) \quad (14)$$

where  $\alpha_i$  is a Lagrange multiplier.

The local minimizers  $z_k^*$  is obtained from:

$$\frac{\partial F_\lambda}{\partial z_k} \Big|_{z_k=z_k^*} = 0. \quad (15)$$

We get the membership updating function with different weight values, which can be expressed as:

$$u_{ki} = \frac{1}{\sum_{j=1}^2 \left( \frac{\lambda \|x_i - z_k\|^2 + (1-\lambda) \|\bar{x}_i - z_k\|^2}{\lambda \|x_i - z_j\|^2 + (1-\lambda) \|\bar{x}_i - z_j\|^2} \right)^{\frac{1}{m-1}}} \quad (16)$$

The original membership updating formula of the standard FCM algorithm is not appropriate to update the membership values of each pixel. By using decomposition mechanism, this paper updates the membership values according to the weight values of each subproblem. The new membership updating formula make full use of the information of the difference image and the filtered difference image.

#### 4.3. Experimental study

In the following experiments, a simulated data set and two real data sets (Bern and Ottawa) with different characteristics are considered to demonstrate the effectiveness of the proposed SAR image change detection method.

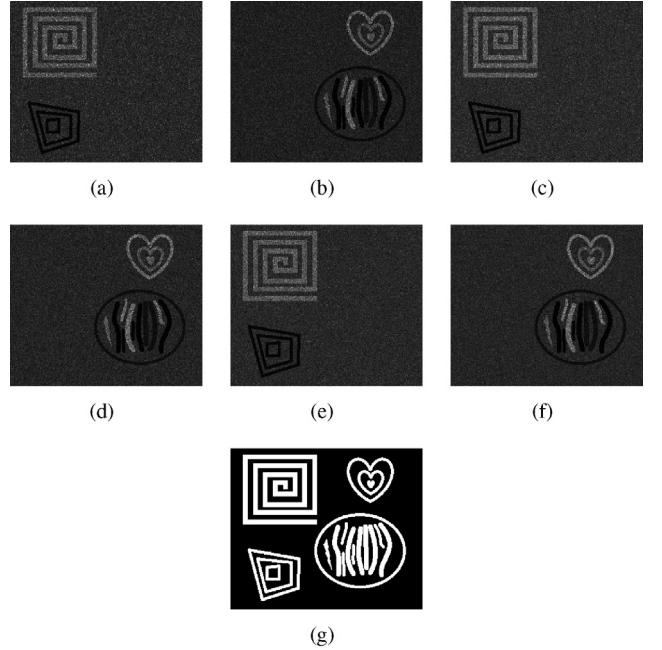


Fig. 12. Simulated data set with three pairs of SAR images with different ENLs. (a) and (b): ENL = 3. (c) and (d): ENL = 4. (e) and (f): ENL = 5. (g) The reference image.

**1) Data sets and quantitative measures:** The first data set is a simulated data set. As shown in Fig. 12, this data set has three pairs of simulated SAR images with ENL = 3, 4, and 5, respectively. And it is used here to analyze the influence of noise level on the change detection results. The reference image is shown in Fig. 12(g).

The second data set that is provided by the Defence Research and Development Canada (DRDC) reveals a section (290 × 350 pixels) of two SAR images acquired by RADARSAT SAR sensor over the city of Ottawa, Canada. The two images mainly present the areas where they were once afflicted with floods in May and August 1997 respectively. The available ground truth is generated by integrating prior information. The images with photo interpretation are shown in Fig. 13.

The third data set that is provided by the European Remote Sensing 2 satellite SAR sensor shows a section (301 × 301 pixels) of two SAR images over an area near the city of Bern.

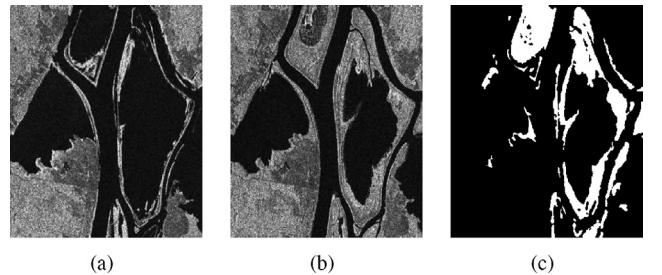


Fig. 13. Ottawa data set: (a) The image acquired in May 1997. (b) The image acquired in August 1997. (c) The ground truth image.

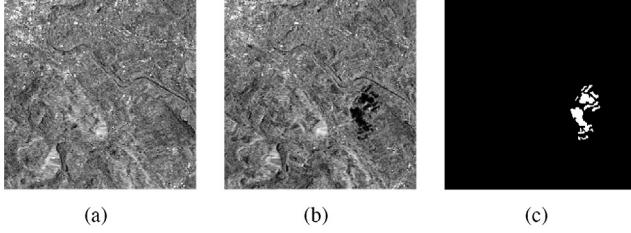


Fig. 14. Bern data set: (a) The image acquired in April 1999. (b) The image acquired in May 1999. (c) The ground truth image.

The river Aare submerged parts of the cities of Thun and Bern and the airport of Bern entirely from April to May 1999. Therefore, we select the Aare valley between Bern and Thun as a test site for detecting changed areas. The available ground truth is generated by integrating prior information. The images with photo interpretation are shown in Fig. 14.

The quantitative analysis of change detection results is set as follow. The first quantitative measure is missed alarms (MAs) which means the number of changed pixels in the change detection result that are incorrectly classified as unchanged when compared to the reference map. These pixels are the missed alarms of the actually changed pixels. The second quantitative measure is false alarms (FAs) which means the number of unchanged pixels in the change detection result that are incorrectly classified as changed when compared to the reference map. These pixels are false detections of the actually unchanged pixels. In order to roughly assess image details and noise of the change detection maps, in the experiments, they are reported as percentages, which are shown as

$$MA(\%) = \frac{MA}{N_c}, FA(\%) = \frac{FA}{N_u} \quad (17)$$

where  $N_c$  and  $N_u$  represent the number of changed and unchanged pixels in the reference image, respectively.

To evaluate the result further, the percentage correct classification (PCC) is used in the experiments, which is defined as

$$PCC = \frac{N - MA - FA}{N} \quad (18)$$

and it expresses the correct rate of the result roughly.  $N$  is the number of the pixels of the difference image. Finally we compute the kappa coefficient which is an overall evaluation criteria used to evaluate the effect of the result in the domain of image segmentation. The higher the value of kappa coefficient is, the better the segmentation result is.

In order to assess the effectiveness of the proposed image change detection approach, we consider the above data sets and compare the results provided by the proposed technique with those obtained by four methods which are widely used in change detection tasks. The generalized Gaussian KI (GGKI) thresholding [41] which selects the global threshold automatically is a very representative threshold algorithm and successfully used in image change detection. The expectation-maximization-based level set method (EMLSM) [48] uses the

expectation-maximization to estimate the mean values of changed and unchanged pixels in the difference image and then adds two energy terms into the level set method to improve accuracy, which is chosen as a method used to compare with the proposed approach. Moreover, the fuzzy C-means algorithm (FCM) (serving as a fundamental one) and the reformulated fuzzy local-information C-means algorithm (RFLICM) [44] are used to compare with the proposed algorithm.

*2) Results on the simulated data set:* As shown in Fig. 15, eight representative maps are chosen from the final maps of the simulated data set with ENL = 4. Obviously, there exists much noise in (a), (b), and (c). But the changed areas in these maps seem more clear than others. By contrast, the background in (g) and (h) is clean and many changed areas are not detected in these maps. The change detection result in the upper left corner in (a) is very close to the reference image. And the change detection result in the lower left corner in (h) is better than others. In Fig. 16, the changed pixels that remain undetected are labeled with red color, and the unchanged pixels wrongly classified as the changed are labeled with yellow color. In other words, the red points roughly represent the lost details and the yellow points are the residual noise. Obviously, the noise points of the selected maps gradually reduce from Fig. 16(a)–(h).

Fig. 17 shows the change detection results obtained by five different algorithms for the simulated data set with ENL = 4. The change detection maps obtained by EMLSM and FCM have much noise. The upper left corner of (a) obtained by GGKI is worse than others. Table 1 presents the change detection results of the simulated data set with different ENLs.

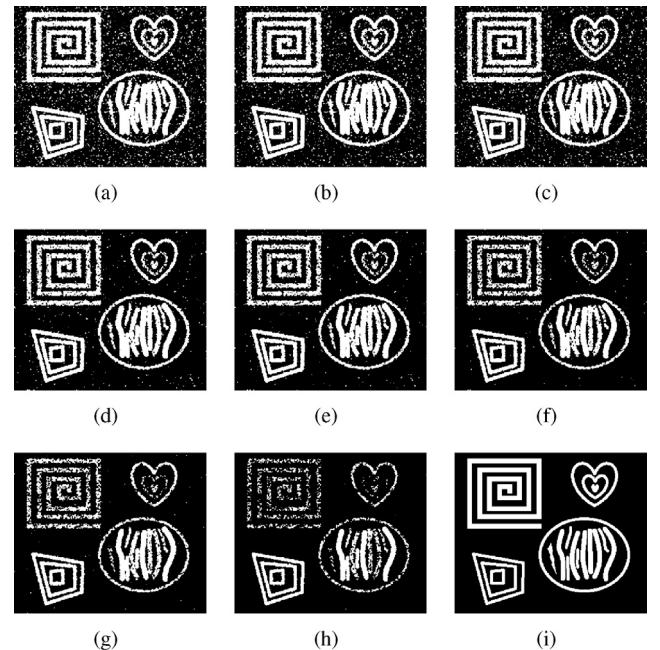


Fig. 15. Change detection results obtained by the proposed method for the simulated data set with ENL = 4. (a–h) Change detection results. (i) The reference image.

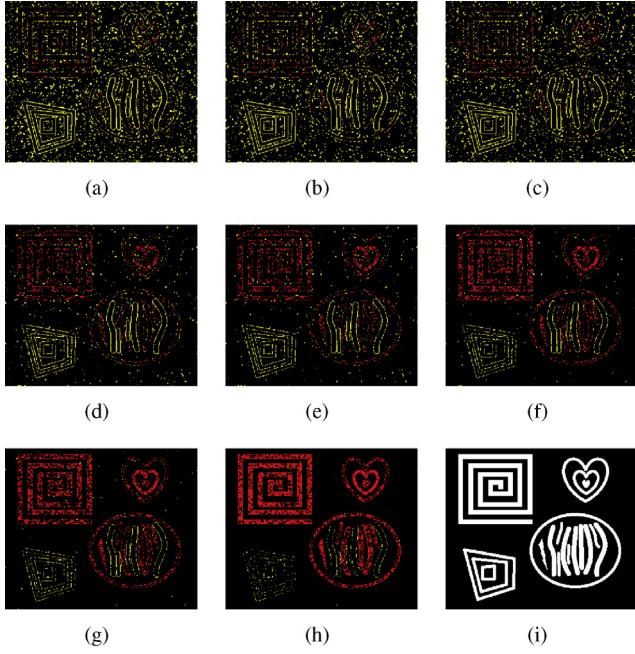


Fig. 16. Comparison of the results shown in Fig. 15 and the reference image of the simulated data set. (a–h) Comparison results. (i) The reference image.

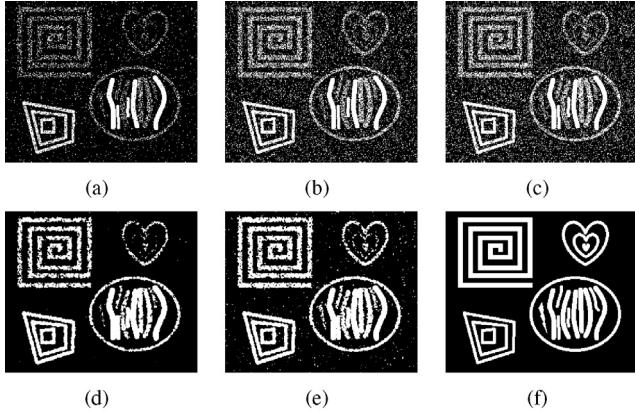


Fig. 17. Change maps of simulated data set with ENL = 4 obtained by (a) GGKI, (b) EMLSM, (c) FCM, (d) RFLICM, and (e) the proposed method. (f) is the reference image.

From Table 1, the GGKI method has less FAs, but has the highest MAs. The RFLICM algorithm has a better performance by incorporating the local information. As described above, a group of change detection maps are obtained by the proposed algorithm. The results with the least total errors are selected to compare with other algorithms, which are shown in Table 1. Obviously, the proposed method has the least TEs and the highest PCCs and Kappa. As shown in Fig. 18, the values of PCC and Kappa obtained by five algorithms gradually increase with the increase of ENL. The proposed method and RFLICM algorithm have higher PCC and Kappa because they can keep better trade-off between preserving details and removing noise. Furthermore, the values of PCC and Kappa obtained by the proposed algorithm are obviously larger than those of RFLICM against the same ENL, respectively.

Table 1

Comparison of the change detection results of simulated data set with different ENLs.

ENL	Method	MA	FA	TE	PCC	Kappa
3	GGKI	12 676	1694	14 370	0.8336	0.4301
	EMLSM	7286	8523	15 809	0.8169	0.5012
	FCM	6473	10 549	17 022	0.8029	0.4873
	RFLICM	5730	1253	6983	0.9191	0.7564
	Proposed	3245	2187	5432	0.9371	0.8281
4	GGKI	11 257	1384	12 641	0.8536	0.5101
	EMLSM	7094	6157	13 251	0.8465	0.5661
	FCM	6308	7755	14 063	0.8371	0.5578
	RFLICM	5065	1220	6285	0.9272	0.7833
	Proposed	2918	2255	5173	0.9401	0.8314
5	GGKI	10 570	1611	12 181	0.8589	0.5336
	EMLSM	7803	4180	11 983	0.8612	0.5885
	FCM	6106	6453	12 559	0.8546	0.5930
	RFLICM	4097	1870	5967	0.9309	0.8001
	Proposed	2189	2538	4727	0.9453	0.8486

3) *Results on the Ottawa data set:* As shown in Fig. 19, eight representative maps are chosen from the final maps of the Ottawa data set. There exists much noise in (a), (b) and (c), and the noise points gradually decrease from Fig. 19(a)–(h). However, the image details (red points shown in Fig. 20) taper off at the same time. The MAs and FAs of these eight maps are shown in Table 2. From Fig. 20 and Table 2, the noise points of the selected maps gradually reduce from (a) to (h). In the classical methods, the final maps often lose some details inevitably, which is similar to Fig. 19(g) or (h). Considering of preserving more details, Fig. 19(c) and (d) seem more useful than others. The results indicate that the proposed method is capable to solve the ill-posed image change detection problem by obtaining a number of different final binary maps in a single run.

Fig. 21 shows the change detection results obtained by the five different algorithms for the Ottawa data set. From intuitive vision of Fig. 21, GGKI, FCM and EMLSM are a little sensitive to noise. The change detection maps obtained by RFLICM and the proposed method show to be closer to the ground truth image which is illustrated in Fig. 21(f). From Table 3, the GGKI method has the most total errors and therefore, the PPC of this method is lowest to 0.9391. We can see obviously that the FLICM and the proposed method have almost the same TEs and PCC. The proposed approach obtains the least TEs 2768 and biggest kappa coefficient 0.8921, which means that the better trade-off between MAs and FAs is achieved.

4) *Results on the Bern data set:* The experiment for the Bern data set is exhibited as well as the previous experiment. Fig. 22(a)–(h) are chosen as eight representative ones from the final maps of the Bern data set. The MAs and FAs of the Bern data set are shown in Table 4. As shown in Fig. 22, there exists much noise in (a), (b) and (c). We can hardly see any noise in (g) and (h). The area A and B are enlarged and shown in Fig. 23, which are marked by the red rectangular in Fig. 22. The first and second row of Fig. 23 correspond to the area A shown in Fig. 22(a)–(h), respectively. The third and fourth

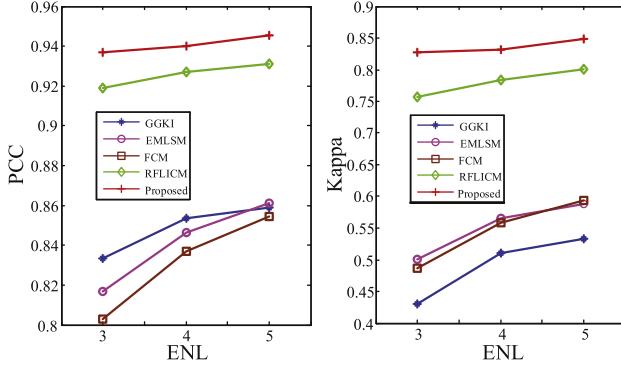


Fig. 18. PCC and Kappa obtained by five different algorithms with different ENLs.

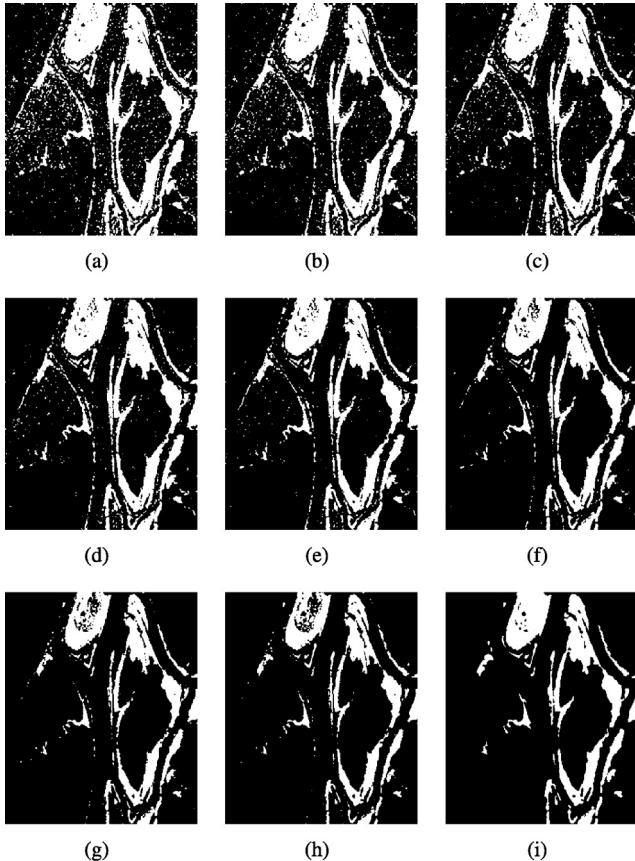


Fig. 19. Change detection results obtained by the proposed method for the Ottawa data set. (a–h) Change detection results. (i) The ground truth image.

row of Fig. 23 are obtained by enlarging the area **B** respectively presented in Fig. 22(a)–(h). From a visual analysis of Fig. 23, we can find that the details in these images gradually reduce from (a) to (h) and (i) to (p) (see Table 4).

Fig. 24 presents a visual comparison between the five different change maps obtained by different methods. The change detection maps obtained by the GGKI and EMLSM method have too much noise. Table 5 shows the change detection results of Bern data set obtained by five different methods. The GGKI method has the highest MAs and FAs and the least PCC and Kappa. The FCM method has the least MAs

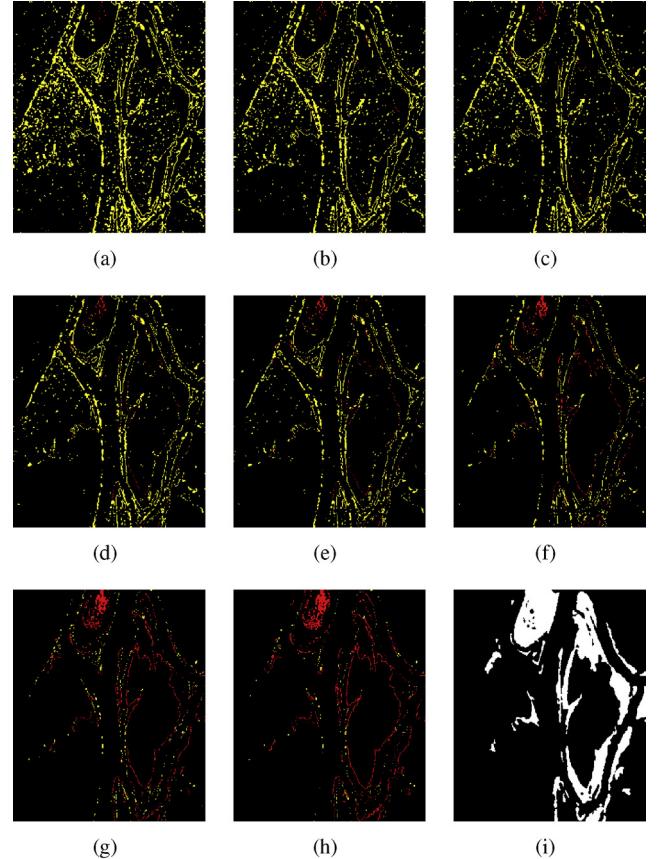


Fig. 20. Comparison of the results shown in Fig. 19 and the ground truth image of the Ottawa data set. (a–h) Comparison results. (i) The ground truth image.

Table 2  
Change detection results on Ottawa data set.

Data set	No.	MA	MA (%)	FA	FA (%)
Ottawa (Fig. 19)	a	61	0.0038	10754	0.1258
	b	117	0.0073	7938	0.0929
	c	145	0.0090	6865	0.0803
	d	296	0.0184	4431	0.0519
	e	421	0.0262	3407	0.0399
	f	784	0.0489	2012	0.0235
	g	1667	0.1039	760	0.0089
	h	2390	0.1489	378	0.0044

but meanwhile gets many FAs. RFLICM and the proposed method have better results than others as they incorporate local information for the purpose of reducing the effect of speckle noise. To sum up, the proposed approach will fit two situations where the changed area appears scattered (the Ottawa data set) and centralized (the Bern data set). For more complicated situation, the proposed method still works well.

## 5. Concluding remarks

In order to solve ill-posed inverse problems in image processing, in this paper, they are modeled as multi-objective optimization problems and optimized by evolutionary algorithms. In the case study of sparse reconstruction, the

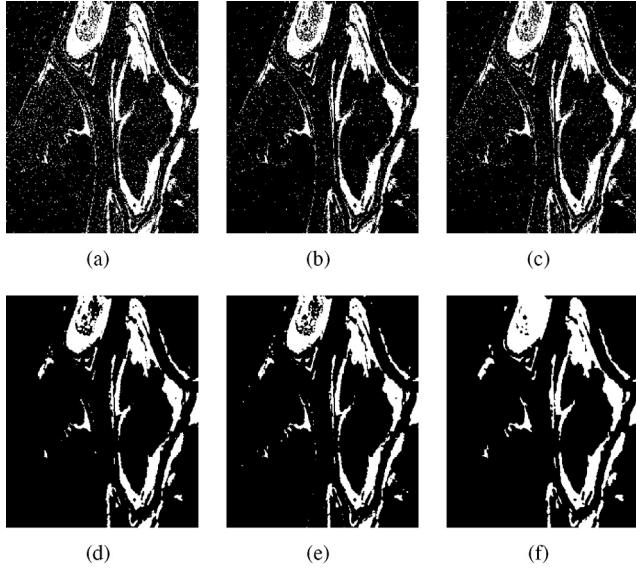


Fig. 21. Change maps of Ottawa data set obtained by (a) GGKI, (b) EMLSM, (c) FCM, (d) RFLICM, and (e) the proposed method. (f) is the ground truth image.

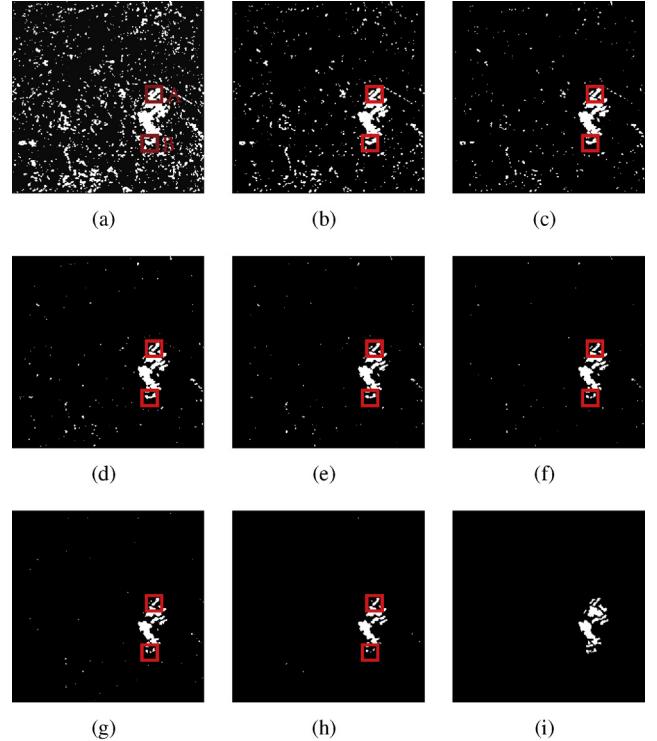


Fig. 22. Change detection results obtained by the proposed method for the Bern data set. (a–h) Change detection results. (i) The ground truth image.

Table 3

Comparison of the change detection results of Ottawa data set with five different methods.

Method	MA	FA	TE	PCC	Kappa
GGKI	2036	4143	6179	0.9391	0.7829
EMLSM	2704	2173	4977	0.9520	0.8171
FCM	2424	2766	5190	0.9489	0.8096
RFLICM	2587	223	2812	0.9723	0.8893
Proposed	2390	378	2768	0.9727	0.8921

Table 4

Change detection results on Bern data set.

Data set	No.	MA	MA (%)	FA	FA (%)
Bern (Fig. 22)	a	7	0.0061	8997	0.1006
	b	12	0.0104	3146	0.0352
	c	29	0.0251	1540	0.0172
	d	46	0.0398	712	0.0080
	f	84	0.0727	430	0.0048
	e	122	0.1056	269	0.0030
	g	167	0.1446	182	0.0020
	h	196	0.1697	101	0.0011

measurement error term and the sparsity term have been optimized by a multi-objective evolutionary algorithm with cooperative coevolutionary. The sparse coefficients are chosen as the decision vector instead of encoding all pixels. In the case study of change detection, two objectives aiming at preserving details and restraining noise have been designed and optimized by the MO framework. The experiments on sparse denoising of natural images, sparse unmixing of hyperspectral data and SAR images change detection demonstrate the effectiveness of MO framework for image ill-posed problems, which can keep the optimal trade-off between multiple conflicting objectives. In the future, we will explore

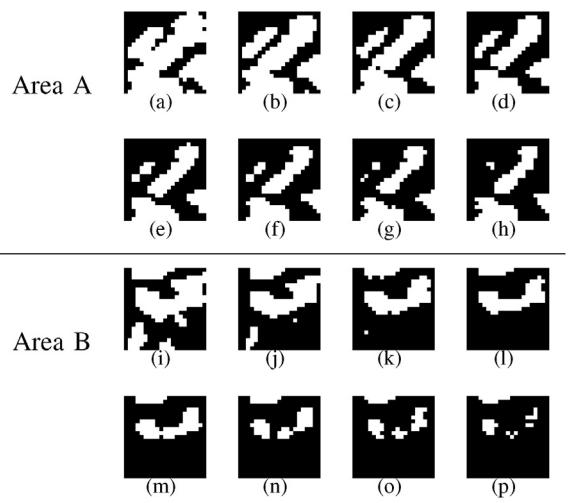


Fig. 23. The area A and B shown in Fig. 22.

Table 5

Comparison of the change detection results of Bern data set with five different methods.

Method	MA	FA	TE	PCC	Kappa
GGKI	301	422	720	0.9921	0.7002
EMLSM	123	216	339	0.9963	0.8570
FCM	94	382	476	0.9947	0.8141
RFLICM	194	122	316	0.9965	0.8656
Proposed	196	101	297	0.9967	0.8643

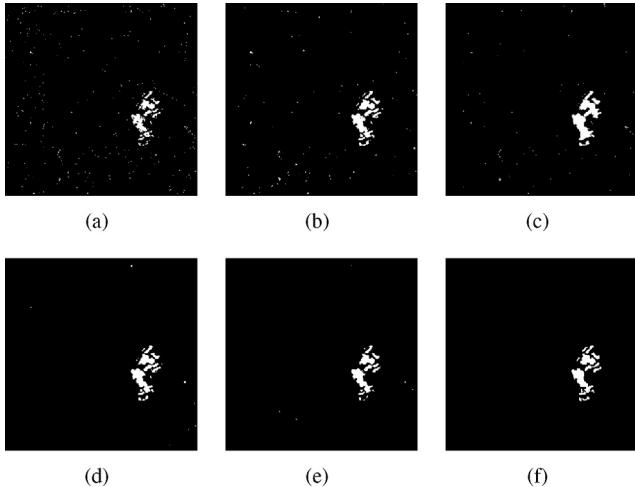


Fig. 24. Change maps of Bern data set obtained by (a) GGKI, (b) EMLSM, (c) FCM, (d) RFLICM, and (e) the proposed method. (f) is the ground truth image.

more objective functions to be suitable for image ill-posed problems using the MO framework and pay interest in improving the search strategy to reduce the time complexity.

## Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant no. 61273317 and 61422209), the National Top Youth Talents Program of China, the Specialized Research Fund for the Doctoral Program of Higher Education (Grant no. 20130203110011) and the Fundamental Research Fund for the Central Universities (Grant no. K5051202053).

## References

- [1] J. Hadamard, Lectures on the Cauchy Problem in Linear Partial Differential Equations, Yale University Press, New Haven, 1923.
- [2] R. Rangarajan, R. Raich, A. Hero, IEEE J. Sel. Top. Signal Process. 1 (1) (2007) 67–78.
- [3] P.-Y. Chen, I. Selesnick, IEEE Trans. Signal Process. 62 (13) (2014) 3464–3478.
- [4] S. Xiang, G. Meng, Y. Wang, C. Pan, C. Zhang, Int. J. Comput. Vis. (2014) 1–24.
- [5] J. Jin, Y. Gu, S. Mei, IEEE J. Sel. Top. Signal Process. 4 (2) (2010) 409–420.
- [6] A.N. Tikhonov, V.Y. Arsenin, Solutions of Ill-posed Problems, Winston and Sons, Washington, DC, 1977.
- [7] L.I. Rudin, S. Osher, E. Fatemi, Phys. D. 60 (1) (1992) 259–268.
- [8] Q.V. Tran, H.T. Nguyen, V.T. Nguyen, D.T. Dang, Appl. Math. Model. 38 (17–18) (2014) 4460–4479.
- [9] V. Vasin, S. George, Appl. Math. Comput. 230 (2014) 406–413.
- [10] C.A.C. Coello, D.A. Van Veldhuizen, G.B. Lamont, Evolutionary Algorithms for Solving Multi-objective Problems, Kluwer, Norwell, MA, 2002.
- [11] C.M. Fonseca, P.J. Fleming, Evol. Comput. 3 (1) (1995) 1–16.
- [12] K. Deb, Multi-objective Optimization Using Evolutionary Algorithms, Wiley, New York, 2001.
- [13] K. Miettinen, Nonlinear Multiobjective Optimization, Kluwer, Norwell, MA, 1999.
- [14] J.D. Schaffer, Multiple objective optimization with vector evaluated genetic algorithms, in: Proc. 1st Int. Conf. Genet. Algorithms, 1985, pp. 93–100.
- [15] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, IEEE Trans. Evol. Comput. 6 (2) (2002) 182–197.
- [16] S. Bandyopadhyay, S. Saha, U. Maulik, K. Deb, IEEE Trans. Evol. Comput. 12 (3) (2008) 269–283.
- [17] Q. Zhang, H. Li, IEEE Trans. Evol. Comput. 11 (6) (2007) 712–731.
- [18] H. Li, Q. Zhang, IEEE Trans. Evol. Comput. 13 (2) (2009) 284–302.
- [19] M.A. Figueiredo, R.D. Nowak, S.J. Wright, IEEE J. Sel. Top. Signal Process. 1 (4) (2007) 586–597.
- [20] D. Needell, R. Vershynin, IEEE J. Sel. Top. Signal Process. 4 (2) (2010) 310–316.
- [21] T. Blumensath, M.E. Davies, IEEE J. Sel. Top. Signal Process. 4 (2) (2010) 298–309.
- [22] Z. Xu, X. Chang, F. Xu, H. Zhang, IEEE Trans. Neural Netw. Learn. Syst. 23 (7) (2012) 1013–1027.
- [23] Z. Yang, K. Tang, X. Yao, Inf. Sci. 178 (15) (2008) 2985–2999.
- [24] Y. Mei, X. Li, X. Yao, IEEE Trans. Evol. Comput. 18 (3) (2014) 435–449.
- [25] M.N. Omidvar, X. Li, Y. Mei, X. Yao, IEEE Trans. Evol. Comput. 18 (3) (2014) 378–393.
- [26] A. Neubauer, A theoretical analysis of the non-uniform mutation operator for the modified genetic algorithm, in: Proc. IEEE Congr. Evol. Comput., 1997, pp. 93–96.
- [27] Z. Michalewicz, Genetic Algorithms + Data Structures = Evolution Programs, Springer-Verlag, 1996.
- [28] L. Li, X. Yao, R. Stolk, M. Gong, S. He, IEEE Trans. Evol. Comput. 18 (6) (2014) 827–845.
- [29] J. Branke, K. Deb, H. Dierolf, M. Osswald, Finding knees in multi-objective optimization, in: Proc. 8th Int. Conf. Parallel Problem Solving from Nature, vol. 3242, 2004, pp. 722–731.
- [30] D. Donoho, Y. Tsaig, IEEE Trans. Inf. Theory 54 (11) (2008) 4789–4812.
- [31] S.J. Wright, R.D. Nowak, M.A. Figueiredo, IEEE Trans. Signal Process. 57 (7) (2009) 2479–2493.
- [32] A. Beck, M. Teboulle, SIAM J. Imag. Sci. 2 (1) (2009) 183–202.
- [33] J. Yang, Y. Zhang, SIAM J. Sci. Comput. 33 (1) (2011) 250–278.
- [34] J.M. Bioucas-Dias, M.A. Figueiredo, Alternating direction algorithms for constrained sparse regression: application to hyperspectral unmixing, in: Proc. 2nd Workshop Hyperspectr. Image Signal Process.: Evol. Remote Sens., 2010, pp. 1–4.
- [35] A.S. Charles, B.A. Olshausen, C.J. Rozell, IEEE J. Sel. Top. Signal Process. 5 (5) (2011) 963–978.
- [36] Y. Zhong, R. Feng, L. Zhang, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (6) (2014) 1889–1909.
- [37] R. Feng, Y. Zhong, L. Zhang, ISPRS J. Photogramm. Remote Sens. 97 (2014) 9–24.
- [38] M.-D. Iordache, J.M. Bioucas-Dias, A. Plaza, IEEE Trans. Geosci. Remote Sens. 50 (11) (2012) 4484–4502.
- [39] F. Bovolo, C. Marin, L. Bruzzone, IEEE Trans. Geosci. Remote Sens. 51 (4) (2013) 2042–2054.
- [40] C. Marin, F. Bovolo, L. Bruzzone, IEEE Trans. Geosci. Remote Sens. 53 (5) (2014) 2664–2682.
- [41] Y. Bazi, L. Bruzzone, F. Melgani, IEEE Trans. Geosci. Remote Sens. 43 (4) (2005) 874–887.
- [42] Y. Ban, O. Yousif, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 5 (4) (2012) 1087–1094.
- [43] H. Hu, Y. Ban, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (8) (2014) 3248–3261.
- [44] M. Gong, Z. Zhou, J. Ma, IEEE Trans. Image Process. 21 (4) (2012) 2141–2151.
- [45] O. Yousif, Y. Ban, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (10) (2014) 4288–4300.
- [46] M. Gong, L. Su, M. Jia, W. Chen, IEEE Trans. Fuzzy Syst. 22 (1) (2014) 98–109.
- [47] O. Yousif, Y. Ban, IEEE Trans. Geosci. Remote Sens. 51 (4) (2013) 2032–2041.
- [48] M. Hao, W. Shi, H. Zhang, C. Li, IEEE Geosci. Remote Sens. Lett. 11 (1) (2014) 210–214.



**Maoguo Gong** (M'07-SM'14) received the B.S. degree in electronic engineering (first class honors) and the Ph.D. degree in electronic science and technology from Xidian University, Xi'an, China, in 2003 and 2009, respectively. Since 2006, he has been a Teacher with Xidian University. In 2008 and 2010, he was promoted as an Associate Professor and as a Full Professor, respectively, both with exception admission. His research interests are in the area of computational intelligence with applications to optimization, learning, data mining and image understanding. Dr. Gong received the prestigious National Program for the support of Top-Notch Young Professionals from the Central Organization Department of China, the Excellent Young Scientist Foundation from the National Natural Science Foundation of China, and the New Century Excellent Talent in University from the Ministry of Education of China. He is the Vice Chair of the IEEE Computational Intelligence Society Task Force on Memetic Computing, an Executive Committee Member of the Chinese Association for Artificial Intelligence, and a Senior Member of the Chinese Computer Federation.



**Hao Li** received the B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 2013. He is currently pursuing the Ph.D. degree in pattern recognition and intelligent systems at the School of Electronic Engineering, Xidian University, Xi'an, China. His research interests include computational intelligence and image understanding.



**Xiangming Jiang** received the B.S. degree in mathematics and applied mathematics from Xidian University, Xi'an, China, in 2015. He is currently pursuing the Ph.D. degree in pattern recognition and intelligent systems at the School of Electronic Engineering, Xidian University, Xi'an, China. His research interests include image understanding and inverse problem.