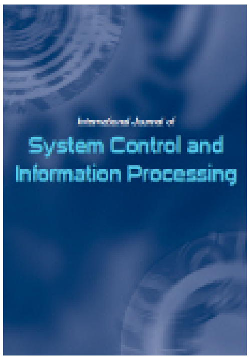


International Journal of System Control and Information Processing

ISSN online: 1759-9342 - ISSN print: 1759-9334
https://www.inderscience.com/ijscip

Sub-pixel resolution thematic map based on multi-source data fusion

Peng Wang, Yuzhou Liu

DOI: 10.1504/IJSCIP.2023.10064386

Article History:

Received:	06 September 2023
Last revised:	08 September 2023
Accepted:	26 March 2024
Published online:	27 June 2025

Sub-pixel resolution thematic map based on multi-source data fusion

Peng Wang* and Yuzhou Liu

Key Laboratory of System Control and Information Processing,
Ministry of Education,
Shanghai, 200030, China

and

Hebei Key Laboratory of Resource and Environmental Disaster
Mechanism and Risk Monitoring,
Institute of Disaster Prevention,
Sanhe, 065201, China

and

College of Electronic and Information Engineering,
Nanjing University of Aeronautics and Astronautics,
Nanjing, 210016, China
Email: w2008084109@126.com
Email: andy704liu@163.com
*Corresponding author

Abstract: In this paper, super-resolution mapping method is used to enhance the spatial resolution of the original hyperspectral images (HSI) and process the mixed pixels, producing the sub-pixel resolution thematic map. Currently, super-resolution mapping usually selects only panchromatic images (PAN) or light detection and ranging images (LiDAR) to be fused with HSI, while this paper proposes a multi-source data fusion method for super-resolution mapping (MSDF), which first fuses HSI with PAN by pansharpening based on principal component analysis (PCA), followed by LiDAR by generalised fusion map method, and finally classifies the fused image by support vector machine (SVM) to generate the sub-pixel resolution thematic map. The proposed MSDF in this paper fuses three types of data sources (i.e., HSI, PAN and LiDAR) to obtain the fused image with full spatial-spectral-elevation information, which enhances the classification accuracy of the sub-pixel resolution thematic map.

Keywords: multi-source data fusion; HIS; hyperspectral images; PAN; panchromatic images; LiDAR images; super-resolution mapping.

Reference to this paper should be made as follows: Wang, P. and Liu, Y. (2025) ‘Sub-pixel resolution thematic map based on multi-source data fusion’, *Int. J. System Control and Information Processing*, Vol. 4, No. 3, pp.204–219.

Biographical notes: Peng Wang received the BE degree in Microelectronics and the PhD degree in Information and Communications Engineering from the College of Information and Communications Engineering, Harbin Engineering University, Harbin, China, in 2012 and 2018, respectively. He is currently an Associate Professor and Doctoral Supervisor with the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research interests include remote-sensing imagery processing and machine learning.

Yuzhou Liu received the BE degree in Electronic Information Engineering from Suzhou University of Science and Technology, in 2023. He is currently pursuing the MS degree in the College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics. His research interests include remote-sensing imagery processing and machine learning.

1 Introduction

Since the 1960s, remote sensing technology has made great progress in theory, technology and applications. Imaging spectrometers on different space platforms are capable of imaging target areas in the mid-infrared, near-infrared, visible and ultraviolet regions of the electromagnetic spectrum in hundreds of successively subdivided spectral bands to obtain hyperspectral images (HSI). HSI have hundreds of spectral bands, which can capture the spectral characteristics of a feature while reflecting its overall morphology and its relationship with the surrounding features, and the rich information it provides can significantly improve the quality, detail and credibility of the subsequent analysis, which has not only received wide attention in remote sensing, but also been put into application in various aspects such as medicine, biology, agriculture, and so on. However, the spectral resolution and spatial resolution of remote sensing images have mutual constraints, and while the spectral resolution is improved, the spatial resolution is also limited. In addition, it is more difficult to acquire wide-area high spatial resolution feature images due to factors such as the high price of the equipment needed to acquire high-precision remote sensing images and the relatively small coverage area of high spatial resolution satellites. Therefore, the use of low-cost equipment to acquire wide-area low spatial resolution remote sensing images, and then processing remote sensing images to obtain high spatial resolution images through technical means has become a research topic with high application value.

In addition, in low spatial resolution remote sensing images, the existence of mixed pixels is unavoidable, i.e., due to the low spatial resolution, the pixels seldom consist of a single homogeneous feature type, which brings uncertainty for information classification and extraction. Processing remote sensing images containing mixed pixels mainly involves solving the following three problems: determining the feature types contained in mixed pixels; calculating the proportion of each feature type; and determining the spatial distribution location of each feature type. End-element extraction selects end-element spectra by semi-automatic or fully-automatic methods to determine the feature types. Mixed pixel unmixing technique can utilise spectral mixing models to calculate the proportion (i.e., abundance) of each feature type in the hybrid image.

However, the end-element extraction and mixed pixel unmixing techniques cannot determine the specific location of each feature type in space, in order to improve the spatial resolution and get the spatial distribution of each feature, Atkinson (1997) formally proposed the theory of super-resolution mapping for the first time. Super-resolution mapping, also known as super-pixel mapping in some literatures, is a technique to improve the spatial resolution of remote sensing images, which not only obtains the proportion of each feature type, but also compensates for the spatial details as much as possible by determining the specific spatial locations of different types of features in the mixed pixels of remote sensing images. It is also able to obtain sub-pixel resolution thematic map with higher spatial resolution by determining the specific spatial locations of different types of features in the mixed image elements of remote sensing images and compensating the spatial detail information as much as possible.

Since super-resolution mapping was first proposed by Atkinson (1997), it has received attention from scientists in the remote sensing community and many other fields, and more and more algorithms have been proposed and rapidly developed. At present, there are mainly the following categories of super-resolution mapping algorithms: the methods based on probability distribution and statistical assumptions, the methods based on optimisation algorithms, and the methods based on supervised neural network models.

The methods based on probability distributions and statistical assumptions refer to the following: assuming that the sub-pixels of a mixed pixel conform to a static statistical distribution, the specific spatial location of the sub-pixels is estimated by determining the maximum probability of their belonging to a category, and the singularities in the pixels that do not conform to this statistical distribution are considered to be in other categories. Typical methods include fuzzy statistical classification, geostatistical indicators, and Markov Random Field models (Kasetkasem et al., 2005). Markov Random Field theory is a theory that analyses spatial or temporal correlation and models an image as a lattice of random variables, where each variable has a dependence on neighbouring random variables other than itself. The model considers the conditional distribution of each pixels with respect to its neighbouring pixels, which can effectively reflect the local distribution characteristics of the pixels. Kasetkasem et al. (2005) from the Agricultural University of Thailand introduced the Markov model into the super-resolution mapping technique (Kasetkasem et al., 2005), and at the same time, the spatial correlation and the spectral information of the mixed image elements were included in the considerations, which reduced the error of the image classification results.

The methods based on optimisation algorithms are one type of the most important super-resolution mapping methods, including linear and nonlinear methods, which uses the spatial structure relationship between pixels to set the specific classification method while satisfying the maximum spatial correlation. Tatem et al. (2001a) first applied the Hopfield neural network (HNN) algorithm to the super-resolution mapping containing two feature types in 2001 (Tatem et al., 2001a), and regarded the sun-pixels in the pixel as neurons in the neural network, and the sub-pixels and their neighbouring sub-pixels as neurons in the neural network, and was successfully applied to a wide range of scenarios, including urban areas (Tatem et al., 2001b) and agricultural applications (Tatem et al., 2001c, 2003). After that, Nyugen et al. utilised the traditional HNN model and added radar images (Nguyen et al., 2005), fused images (Nguyen et al., 2006), and panchromatic band (Nguyen et al., 2011) as auxiliary data, respectively, to obtain accurate mapping results. Ling et al. (2010) proposed a new super-resolution mapping method by using HNN as the basis of multiple sub-pixels captured by satellites. In addition, Atkinson proposed the classical pixel swapping algorithm (PSA) (Atkinson et al., 2005), this algorithm is applicable to remote sensing images containing two classes of features, firstly, initialise an image randomly, and exchange the category corresponding to the sub-pixel with the smallest spatial correlation within each pixel, so that the spatial correlation becomes larger and larger by iteration, and the stable ideal image is obtained. Mertens et al. (2006) proposed a super-resolution mapping algorithm based on the pixel/pixel spatial attraction model (SPSAM), and Shen et al. (2009) used the mapping results of SPSAM as the initial image of the classical PSA algorithm. The initial image of classical PSA algorithm can make up for the shortcomings of classical PSA algorithm which uses pseudo-random numbers for initialisation, and improve the image resolution and stability of super-resolution mapping. Ling et al (2005) applied cellular automate (CA) to super-resolution mapping, and interchanged the positions of cells to increase the image resolution between image pixels.

The methods based on supervised neural network models are well suited to simulate complex structural relationships, and the neural network simulates sub-pixel mapping results by training pixels that are adjacent to the target type and adjusting the network parameters. The back propagation neural network (BPNN), as an alternative neural network model besides the HNN, has also been applied to super-resolution mapping (Mertens et al., 2003). BPNN firstly created a model by learning the feature distribution information of high-resolution images and training to get the weight parameters in the neural network, and then input the low-resolution images into the model to get the super-resolution mapping results; Zhang et al. (2008) proposed the BPMAP algorithm by combining the BPNN with the observation model; Xu et al. (2011) used the spatial autocorrelation function Moran'I to constrain the super-resolution mapping results of BPNN, which improves the applicability of the model in the sub-pixel space. In addition, Boucher and Kyriakidis (2006) proposed a super-resolution mapping algorithm based on geostatistical differences, which utilises the cooperative kriging method to obtain the probability of occurrence of various types of sub-pixels, and to obtain high-resolution mapping results. This method applies geostatistics to super-resolution mapping without iteration and with high computational efficiency, but it has limitations in dealing with linear features.

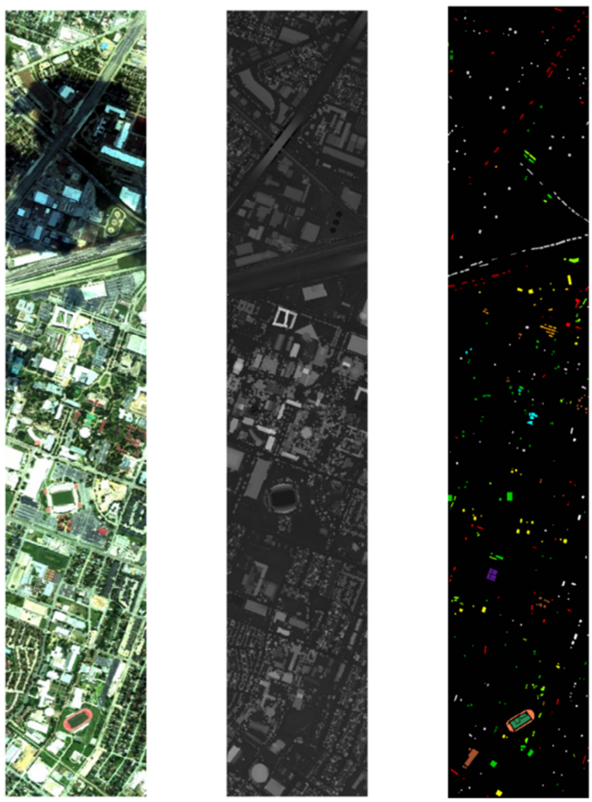
However, through the analysis of these literatures, we can find that the type of auxiliary data used in super-resolution mapping is relatively single. Multiple types of auxiliary information are not effectively used at the same time in the existing super-resolution mapping, resulting in hindering the further improvement of accuracy. In order to solve this problem, this paper proposes a multi-source data fusion method for super-resolution mapping (MSDF), which first fuses HSI with PAN by pansharpening based on principal component analysis (PCA), followed by light detection and ranging images (LiDAR) by generalised fusion map method, and finally classifies the fused image by support vector machine (SVM) to generate the sub-pixel resolution thematic map. The contributions of this work are as follows:

1. Full spatial-spectral- elevation information is utilised by fusing HSI, PAN, and LiDAR image.
2. Under the condition of multi-source data, it provides a more effective ideal for super-resolution mapping technology.
3. Experimental results show that the proposed MSDF is superior to the state-of-the-art super-resolution mapping methods.

2 Study data

We choose 2013 IEEE GRSS Data Fusion Contest as the study data. The datasets are obtained by the NSF-funded Center for Airborne Laser Mapping (NCALM) at the University of Houston campus and its adjacent downtown area in June 2012 for the simulation validation. The datasets include the HSI, LiDAR and classification reference image. As shown in Figure 1(a), there are 349×1905 pixels and 144 bands in 380–1050 nm spectral resolution in HSI. Figure 1(b) shows the LiDAR with 349 × 1905 pixels and 2.5 m spatial resolution. The classification reference image includes 15 classes shown in Figure 1(c).

Figure 1 Study data: (a) false colour image of HSI; (b) LiDAR and (c) reference image (see online version for colours)



(a) (b) (c)

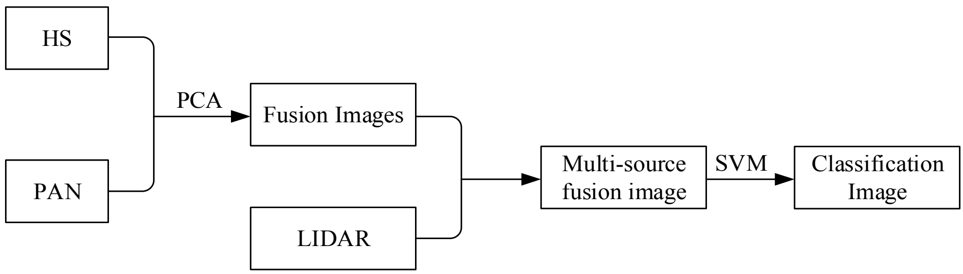
3 The proposed method

Aiming at the problem of super-resolution mapping for multi-source data fusion, the flowchart of the proposed MSDF is shown in Figure 2, which mainly accomplishes the following:

1. We use pansharpening based on the PCA method to fuse the HSI and PAN, replace the spatial structure information in the low spatial resolution HSI with the high spatial resolution PAN, producing the high spatial resolution fusion image;
2. The fusion image and the LiDAR are fused by feature fusion to produce the multi-source fusion image including spatial-spectral-elevation information.

3 Use the support vector machine (SVM) with radial basis function as the kernel function to classify the multi-source fusion image to obtain the sub-pixel resolution thematic map.

Figure 2 Flowchart of the proposed MSDF



3.1 Pansharpening

3.1.1 Ingredient substitution principle

Principal component analysis (PCA) is one of the panchromatic sharpening methods based on component substitution (CS). The fusion steps of Component Substitution-based panchromatic sharpening are generally as follows: first, in order to separate the spatial structure information from the spectral information, CS projects the original low spatial resolution hyperspectral image to another space; then, the spatial structure information of the spatial resolution hyperspectral image is replaced with the high spatial resolution panchromatic image, and at this time, the spatial correlation between the panchromatic image and the replaced low spatial resolution spatial structure information is as follows: the larger the spatial resolution, the smaller the spectral distortion. The larger the spatial correlation between the panchromatic image and the substituted low spatial resolution spatial structural information, the smaller the spectral distortion, so before this step of substitution, the panchromatic image can be histogram-matched with the substituted component to show the same mean and variance between the two; and finally, the final fused image is obtained by inverse transformation. The model of this method can be represented as:

$$\widehat{Y^b} = \widetilde{Y^b} + g_b\left(P - I\right) \tag{1}$$

where b denotes the bth spectral band, $b = 1, 2, \ldots, N$, N is the total number of hyperspectral image bands, Y is the original low spatial resolution hyperspectral image, $\widehat{Y^b}$ is the image after panchromatic sharpening of the bth spectral band, $\widetilde{Y^b}$ is the image obtained by interpolating the original low-resolution image of the bth band to the size of the panchromatic image, $g_b = \left[g_1, g_2, \ldots, g_N\right]$ is the gain vector, P is the high spatial resolution panchromatic image, and I may be expressed as:

$$I = \sum_{b=1}^{N} w_b \widetilde{Y^b} \tag{2}$$

where $w_b = \left[w_1, w_2, \ldots, w_N\right]^T$ is the weight vector, which is used to measure the spectral overlap between spectral bands and panchromatic images (PAN).

3.1.2 Principles of principal component analysis

In this paper, we choose to use the PCA algorithm in CS to fuse the HS image and PAN image to improve the spatial resolution of the original hyperspectral image. The PCA method can simplify the processing of multidimensional information, and represent the hyperspectral image with a few components that contain rich information, the components are independent of each other, and the processing of a single component has less influence on other components, so that we can avoid the appearance of redundant information as much as possible. The PCA method's The main idea is to perform a component transformation on the HS image to obtain several principal components, then the first principal component is substituted with the PAN image for the components, and the final fused image is obtained using inverse transformation. The PCA method is described in mathematical language as follows:

1 Calculate the covariance matrix C. Represent each band of the hyperspectral image as a vector form as:

$$X = \left[x_1, x_2, \ldots, x_n\right]^T \tag{3}$$

The variance between different bands can be expressed as:

$$\sigma_{i,j}^2 = E\left[\left(x_i - u_i\right)\left(x_j - u_j\right)\right], \left(i, j = 1, 2, \ldots, n\right) \tag{4}$$

Then the covariance matrix can be derived as:

$$C = \begin{bmatrix} \sigma_{11} & \sigma_{21} & \cdots & \sigma_{n1} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{n2} \\ \vdots & & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_{nn} \end{bmatrix} \tag{5}$$

2 Solve for the eigenvalues and eigenvectors of C according to the following equation:

$$\phi_n = \left[\phi_1, \phi_2, \ldots, \phi_n\right]^T \tag{6}$$

3 Decompose the principal component matrix according to the following equation:

$$Y = \left[\phi_1, \phi_2, \ldots, \phi_n\right] * X \tag{7}$$

It's available:

$$Y = \left[PC_1, PC_2, \ldots, PC_n\right]^T \tag{8}$$

where PC is the principal component and PC_1 is the first principal component.

3.2 LiDAR image fusion method

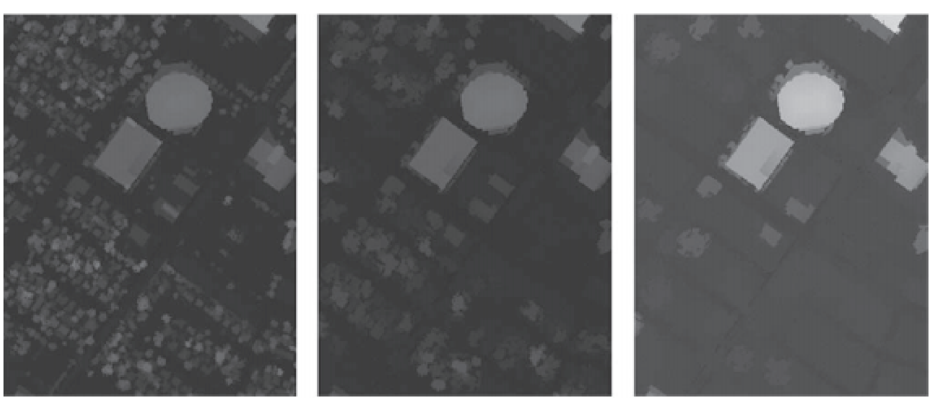
LiDAR is a product of the birth of laser technology, the perfect combination of modern laser technology and traditional radar technology. Since the 21st century, LiDAR technology has been developing rapidly, from the coherent outlier two-dimensional scanning mechanism to the flash focal plane direct detection mechanism, which is able to carry out three-dimensional distance imaging and intensity imaging of the features, and has a very wide range of applications in various fields. LiDAR images can provide height information of the investigated area, which can further enhance the accuracy of spatial structure information and also lay the foundation for subsequent feature classification.

3.2.1 Morphological characteristics

The results generated by reconstructing an image by using structural elements (SEs) of a particular size and shape to perform morphological open and close operations are called morphological profiles. For example, a morphological profile (MP) using a disk-shaped SE can reflect the minimum size of the object in question, while an oriented MP shows the maximum size of the processed object (Bellens et al., 2008). When processing LiDAR images, the open operation acts on high-elevation regions measured by LiDAR data, such as the highest part of the roof, while the closed operation acts on low-elevation regions in the LiDAR data, i.e., the open operation removes features whose brightness is less than the SE. By incrementally increasing the SE and repeating this operation, a complete MP can be obtained to provide information such as the size and shape of features in the image.

In the simulation experiments, the LiDAR data and the first two principal components (PCs) of the fused images of HS and PAN, which account for more than 99% of the cumulative variance, are partially reconstructed, and then morphological open and close operations are performed to obtain the results as morphological features. When using the disk-shaped SE, in order to compute the first two PCs for the LiDAR data and the fused images, MP needs to perform 15 open and close operations (from 1 to 15 with a step size of 1). For the linear SE, the maximum value (for the open operation) or minimum value (for the closed operation) needs to be taken every 10° of rotation and partial reconstruction is performed using 10% of the SE length. Then, to compute the LiDAR data and the first two PCs of the fused images, MP performs 20 open and closed operations (from 5 to 100 in steps of 5). Figure 3 shows the MP results for partial reconstruction of LiDAR data with different scales. The bright objects (i.e., high altitude objects) gradually disappear as the size of the open and close operation SE increases.

Figure 3 Reconstruction results for a portion of LiDAR data at progressively larger disc-shaped SE radius sizes (1, 3, and 5)



3.2.2 Integration methods

Before fusing the LiDAR images with the fused images, it is necessary to use the kernel principal component analysis (KPCA) (Bellens et al., 2008) method to normalise the feature dimensions and reduce the noise in the space. In this paper, we choose to standardise the feature dimension to the smallest dimension of the feature source, $D = 70$, and set $X^{Spe} = \left\{X_i^{Spe}\right\}_{i=1}^{n}$ as the spectral feature, $X^{Spa} = \left\{X_i^{Spa}\right\}_{i=1}^{n}$ as the spatial feature of the fused image, and $X^{Ele} = \left\{X_i^{Ele}\right\}_i^n = 1$ as the height feature. After standardisation, we can get $X_i^{Spe} \in R^D$, $X_i^{Spa} \in R^D$ and $X_i^{Ele} \in R^D$. Let $X^{Sta} = \left\{X_i^{Sta}\right\}_{i=1}^{n} = \left[X^{Spe}; X^{Spa}; X^{Ele}\right]$, then $X_i^{Sta} = [\, X_i^{Spe}; X_i^{Spa}; X_i^{Ele} \,] \in R^{3D}$ is the vector consisting of spectral, spatial and height features, space and height features as vectors. Also let $\left\{z_i\right\}_{i=1}^{n}$, $z_i \in R^d$ denote the fused features in the low-dimensional feature space, where $d \leq 3\text{D}$.

In order to combine dimensionality reduction with feature fusion, it is necessary to find a transformation matrix *W*, which should not only fuse different features in the low-dimensional space, but also retain the local neighbourhood information and can detect the manifolds embedded in the high-dimensional feature space. The method for finding the transformation matrix W (Liao et al., 2012a) can be expressed as follows:

$$\arg\min_{W \in R^{3D \times d}} \left(\sum_{i,j=1}^{n} \| W^T x_i - W^T x_j \|^2 A_{ij} \right) \tag{9}$$

where the matrix A is the edge of the graph $G = (X, A)$.

Define the fusion map $G^{Fus} = \left(X^{Sta}, A^{Fus}\right)$ to be

$$A^{Fus} = X^{Spe} \odot X^{Spa} \odot X^{Ele} \tag{10}$$

where "$\odot$" denotes element-by-element multiplication, i.e., $A_{i,j}^{Fus} = A_{i,j}^{Spe} A_{i,j}^{Spa} A_{i,j}^{Ele}$ with $A_{i,j}^{Fus} = 1$ when and only when $A_{ij}^{Spe} = 1$, $A_{ij}^{Spa} = 1$, $A_{ij}^{Ele} = 1$.

Because the above fusion map definition ignores the subtle differences in spectral, altitude, and other information in real situations, which are sometimes not applicable, this paper adopts the generalised fusion map $G^{GFus} = \left(X^{Sta}, Q^{GFus}\right)$. Assuming that Δ is the pairwise distance matrix of the stacked feature X^{Sta}, a fusion distance matrix is defined as

$$\Delta^{GFus} = \Delta + A^{Neg} \max \Delta \tag{11}$$

where $A^{Neg} = \neg A^{Fus}$ and the operator "¬" denotes logical inversion. Let N_i denote the k-NN of x_i for each node x_i ($i \in \{1, \ldots, n\}$), its k-NN, i.e., N_i, is first found in the fusion distance matrix Δ^{GFus}, and then its edge is formally defined as

$$Q_{ij}^{GFus} = \begin{cases} -\| x_i - x_j \|, & \text{if } x_i \in N_i \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

The use of neighbourhood N_i ensures that data points with large differences in any of the spectral, spatial, or altitude features are unlikely to be located within k-NN of other data points, i.e., they will not meet in the fusion map and cause indistinguishability.

To avoid parsimony, the following constraints are used (Liao et al., 2012b):

$$W^T\left(X^{Sta}\right)D^{GFus}\left(X^{Sta}\right)^T W = I \tag{13}$$

where D^{GFus} is the diagonal matrix, $D_{i,i}^{GFus} = \sum_{j=1}^{n} Q_{i,j}^{GFus}$, and I is the unit matrix. The transformation matrix $W = (w1, w2, \ldots, wr)$, obtained by summing up the r eigenvectors associated with the smallest eigenvalue $\lambda 1 \le \lambda 2 \le \cdots \le \lambda r$ of the generalised eigenvalue problem:

$$\left(X^{Sta}\right)L^{GFus}\left(X^{Sta}\right)^T w = \lambda\left(X^{Sta}\right)D^{GFus}\left(X^{Sta}\right)^T w \tag{14}$$

where $L^{GFus} = D^{GFus} - Q^{GFus}$ is the fused Laplace matrix.

3.3 *Classification method*

A SVM is a generalised linear classifier that classifies data binary in a supervised learning manner, using a kernel function to map the sample set vectors to a high-dimensional feature space and forming a constantly moving hyperplane in this high-dimensional space to classify the sample set until different types of sample points lie on both sides of the hyperplane until different types of sample points are located on both sides of the hyperplane. There may be more than one hyperplane that meets the classification requirements of the sample set in the classification result, and the hyperplane that maximises the distance between the two types of samples on both sides is called the optimal decision hyperplane, which has a better generalisation ability when dealing with classification problems.

Assuming that the sample set is $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\}$, the hyperplane that separates the sample set in the high-dimensional feature space can be represented as:

$$w^T x + b = 0 \tag{15}$$

where w is the normal vector of the hyperplane and b is the distance between the hyperplane and the coordinate origin. Equal scaling of w and b can find the optimal decision hyperplane, then the basic type of SVM is:

$$\max_{w,b} \frac{2}{\| w \|}$$

$$s.t.\ y_i\left(w^T x + b\right) \ge 1, \quad i = 1, 2, \ldots, m \tag{16}$$

SVM kernel functions include linear, polynomial, radial basis (RBF) and Sigmoid kernel functions. Compared with other kernel functions, the RBF function has strong localisation and can map the samples to a high-dimensional space. Because the RBF kernel function has fewer parameters compared to the polynomial kernel function, its computational complexity does not change with the change of parameters, and it can get better processing results regardless of the size of the samples, and it satisfies the Mercer condition in all parameter spaces, and it is the most commonly used of all the kernel functions, therefore, in this paper, we choose to use the RBF function as the kernel of the SVM. The RBF kernel function can be expressed as:

$$K(x_t, y_t) = \exp\left(-\frac{\| x_t - y_t \|^2}{2\sigma_t^2}\right) \tag{17}$$

where x_t is the input sample, y_t is the center of the Gaussian function, and σ_t is the variance of the Gaussian function.

The SVM with RBF as the kernel function requires the determination of two parameters: the penalty coefficient *C* and gamma a grid search for *C* and gamma using five-fold cross-validation finds the most appropriate penalty coefficient *C* in $\{10^{-1}, 10^0, 10^1, 10^2, 10^3\}$ and in $\{10^{-3}, 10^{-2}, 10^{-1}, 10^0, 10^1\}$ to find the ideal gamma.

4 Experiment

4.1 Experimental set up

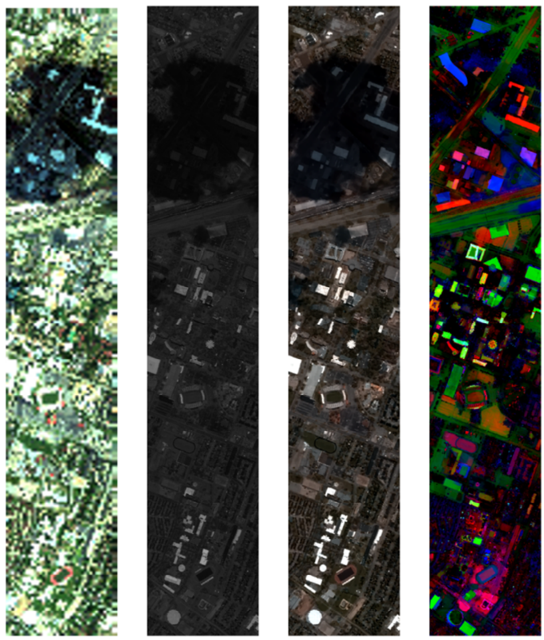
According to the general experimental process of super-resolution mapping, to know the classes at sub-pixel level, we downsampled the HSI in Figure 1(a) with 8×8 mean filter to produce the simulated low resolution HSI, as shown in Figure 4(a). In addition, we only consider pansharpening technique and avoid the influence of errors caused by the acquisition of PAN. As shown in Figure 4(b), the appropriate PAN is obtained by applying the spectral response function of IKONOS satellite to the hyperspectral image. We fused Figure 4(a) and (b) by pansharpening based on PCA to produce the fused image shown in Figure 4(c). We then fuse Figure 4(c) with Figure 1(b) to obtain the multi-source fusion image in Figure 4(d). When compared Figure 4(c) with Figure 4(d), it is noted due to the full spatial-spectral-elevation information, the boundaries and heights among the classes are clearer in Figure 4(d) that these in Figure 4(c). Therefore, when using Figure 4(d), the proposed MRSFD could produce the better mapping result.

Six super-resolution mapping based on fusion methods are tested and compared: Intensity-hue-saturation (IHS), Brovey transform (Brovey), High-pass filtering (HPF), Smoothing filter based on intensity modulation (SFIM), PCA, and the proposed MSDF (Wang et al., 2019). The mapping results from the six methods are evaluated by overall accuracy (OA) and kappa coefficient. All experiments are processed in MATLAB R2018 software package in a Pentium(R) Dual-core Processor (2.20 GHz).

4.2 Analysis of experimental results

Figure 5 shows the sub-pixel resolution thematic maps from six super-resolution mapping methods. Table 1 shows the classification accuracy of each method, using the average accuracy AA, overall accuracy OA, and Kappa coefficient to measure the accuracy of the classification results, in addition to the time required for each method to complete the fusion and classification, listing the classification accuracy for each feature type. As shown in Figure 5(a)–(e), HIS, Brovey, HPF, SFIM and PCA are all belong to the super-resolution methods based on HSI and PAN fusion. As shown in Figure 5(d), it can be seen that the SFIM method handles the feature classification around the ballpark poorly and is not effective in reducing the negative impact of cloud cover on feature classification compared to the other methods. Regardless of the perspective, the SFIM method has the worst time consumption, image fusion results and classification accuracy. The Brovey method is better in treating soil and land type of features, but it takes longer than the PCA method. The PCA method is better and least time consuming in dealing with man-made features such as billboards, commercial buildings, and urban roads, and is second only to the Brovey method in terms of AA, OA, and Kappa coefficients. Therefore, we select the PCA as the HSI and PAN fusion in the proposed MSDF.

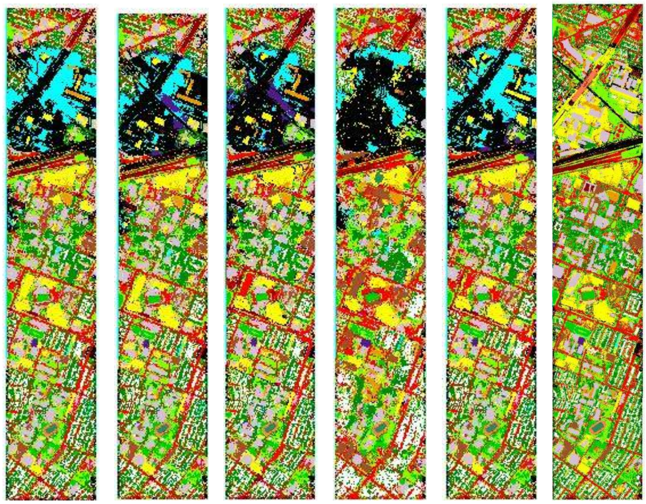
Figure 4 Fusion result: (a) low resolution HSI; (b) PAN; (c) pansharpening result and (d) multi-source fusion result (see online version for colours)



(a) (b) (c) (d)

As can be seen from the above charts, for HIS, Brovey, HPF, SFIM and PCA, the processing results obtained by using only HSI fused with PAN cannot effectively deal with features under cloud cover, so the proposed MSDF solves the problems posed by cloud cover by utilising the elevation information carried by the LiDAR. From Figure 5(f), it can be seen that the fusion of LiDAR can greatly solve the problem of cloud cover. From Table 1, it can be seen that, in addition to some of the feature types, the accuracy of the ground has decreased by a very small margin, and the classification accuracy of many feature types that were originally classified with very low accuracy has increased by a large margin in the proposed MSDF. For example, when compared with PCA, the classification accuracy of soft grass has increased by 20.49%, and the classification accuracy of billboards and commercial buildings has increased by 56.13%. Overall, the OA, AA and Kappa coefficients also improved by 7.81%, 5.27% and 0.08 respectively.

Figure 5 Sub-pixel resolution thematic map: (a) IHS; (b) Brovey; (c) HPF; (d) SFIM; (e) PCA and (f) the proposed MSDF (see online version for colours)



(a) (b) (c) (d) (e) (f)

Table 1 Evaluating indicator of six super-resolution mapping methods (see online version for colours)

	IHS	*Brovey*	*HPF*	*SFIM*	*PCA*	*Proposed MSDF*
OA (%)	81.30	82.32	81.76	70.91	82.18	89.99
AA (%)	84.35	85.20	84.62	76.12	84.98	90.25
Kappa	0.7980	0.8085	0.8025	0.6855	0.8073	0.8914
Time-consuming (s)	265.240	265.042	289.070	320.869	233.820	312.443
Green grass	77.97	79.96	76.45	68.85	79.87	82.15
Soft grass	80.73	77.73	78.10	67.58	78.10	98.59
Artificial turf	100.00	100.00	100.00	99.60	100.00	99.60
Trees	86.17	85.89	82.86	55.97	84.56	98.77
Soil	99.53	100.00	98.48	89.39	99.91	100.00
Rivers, lakes, etc.	99.30	99.30	99.30	93.01	99.30	95.10
Residential building	84.51	86.94	87.22	85.35	86.85	92.44
Billboards and commercial buildings	33.52	36.85	35.42	21.56	39.03	95.16
City road	58.49	89.52	89.24	79.13	89.52	89.05
Freeway	60.81	56.37	62.93	57.24	59.17	51.83
Railroad tracks	90.04	94.97	95.07	79.89	92.22	93.36
Parking lot 1	93.18	92.89	89.82	64.55	91.07	92.41
Parking lot 2	76.14	77.54	74.39	80.00	75.09	65.26
Swimming pools	100.00	100.00	100.00	100.00	100.00	100.00
Runway	100.00	100.00	100.00	99.79	100.00	100.00

5 Conclusion

HSI carry hundreds of spectral bands, which can provide very rich spectral information, and are therefore very widely used in various fields. However, while the spectral resolution increases, the spatial resolution decreases. And the large number of mixed pixels carried in HSI also hinders the clear identification of object types everywhere. Super-resolution mapping technology has been proposed to solve the problem of mixed pixels. However, the type of auxiliary data used in super-resolution mapping is relatively single. Multiple types of auxiliary information are not effectively used at the same time in the existing super-resolution mapping, resulting in hindering the further improvement of accuracy. To fully use spatial-spectral- elevation information, this paper proposes a multi-source data fusion MSDF. In the proposed MSDF, we first improve the spatial resolution by using the PCA fusion of HS and PAN, and we then add the LiDAR to the fusion result of the previous step, finally we classify the multi-source fusion result by using SVM to clearly extract the types of local features which greatly improves the classification accuracy. The experiments on study data show that the proposed MSDF can obtain the best mapping results with the highest mapping accuracy OA (%) of 89.99, AA (%) of 90.25, and Kappa of 0.8914.

Acknowledgement

This work is supported by Hebei Key Laboratory of Resource and Environmental Disaster Mechanism and Risk Monitoring (Grant No. FZ248201), the funding Fundamental Research Funds for the Central Universities in Nanjing University of Aeronautics and Astronautics (Grant No. NS2023020, NJ2023029 and QZJC20230206), Key Laboratory of System Control and Information Processing, Ministry of Education (Grant No. Scip20230104), National Natural Science Foundation of Jiangsu Province (Grant No. BK20221478), Youth Promotion Talent Project of Jiangsu Association for Science and Technology (Grant No. TJ-2023-010), the Open Fund of Hunan Geological Disaster Monitoring Early Warning and Emergency Rescue Engineering Technology Research Center (Grant No. hndzgczx202407).

References

Atkinson, P.M. (2005) 'Sub-pixel target mapping from soft-classified, remotely sensed imagery', *Photogrammetric Engineering and Remote Sensing*, Vol. 71, No. 7, pp.839–846.

Atkinson, P.M. (1997) *Mapping Sub–Pixel Boundaries From Remotely Sensed Images*, Innovmions in GIS IV, Taylor & Francis, London, pp.167–180.

Belkin, M. and Niyogi, P. (2001) 'Laplacian eigenmaps and spectral techniques for embedding and clustering', *Advances in Neural Information Processing Systems 14*, MIT Press, Cambridge, MA, USA, pp.585–591.

Bellens, R., Gautama, S., Martinez-Fonte, L., Philips, W., Chan, C.W. and Canters, F. (2008) 'Improved classification of VHR images of urban areas using directional morphological profiles', *IEEE Trans. Geosci. RemoteSens.*, Vol. 46, No. 10, October, pp.2803–2812.

Boucher, A. and Kyriakidis, P.C. (2006) 'Super-resolution land cover mapping with indicator geostatistics', *Remote Sensing of Environment*, Vol. 104, No. 3, pp.264–282.

Kasetkasem, T., Arora, M.K. and Varshney, P.K. (2005) 'Super-resolution land-cover mapping using a Markov random field based approach', *Remote Sensing of Environment*, Vol. 96, Nos. 3–4, pp.302–314.

Keshava, N. and Mustard, J.F. (2002) 'Spectral unmixing', *IEEE Signal Processing Magazine*, Vol. 19, No. 1, pp.44–57.

Li, X., Cao, C. and Chang, C. (2007) 'The first law of geography and spatial temporal proximity', *Chinese Journal of Nature*, Vol. 29, No. 2, pp.69–71.

Liao, W., Bellens, R., Pižurica, A., Philips, W. and Pi, Y. (2012a) 'Classiffification of hyperspectral data over urban areas using directional morphological profiles and semi-supervised feature extraction', *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, Vol. 5, No. 4, August, pp.1177–1190.

Liao, W., Bellens, R., Pižurica, A., Philips, W. and Pi, Y. (2012b) 'Classiffification of hyperspectral data over urban areas based on extended morphological profile with partial reconstruction', *Proc. ACIVS, Brno, Czech Republic*, pp.278–289.

Ling, F., Zhang, Q., Wang, C. and Zhou, J. (2005) 'Sub-image element localization of remote sensing images based on metacellular automata model', *Chinese Journal of Image Graphics*, Vol. 10, No. 7, p.916–921.

Ling, F., Du, Y., Xiao, F., Xue, H.P. and Wu, S.J. (2010) 'Super-resolution land-cover mapping using multiple sub-pixel shifted remotely sensed images', *International Journal of Remote Sensing*, Vol. 31, No. 19, pp.5023–5040.

Mertens, K.C., De Baets, B. and Wulf, R.R.D. (2003) 'Sub·pixel mapping with neural networks: real-world spatial configurations learned from artificial shapes', *Proceeding of 4th International Symposium on Remote Sensing of Urban Areas*, Regensburg, Germany, pp.117–121.

Mertens, K.C., Basets, B.D., Verbeke, L.P.C. and De Wulf, R. (2006) 'A sub-pixel mapping algorithm based on sub–pixel/pixel spatial attraction models', *International Journal of Remote Sensing*, Vol. 27, No. 15, pp.3293–3310.

Nguyen, M.Q., Atkinson, P.M. and Lewis, H.G. (2005) 'Superresolution mapping using a Hopfield neural network with LiDAR data', *IEEE Geoscience and. Remote Sensing Letters*, Vol. 2, No. 3, pp.366–370.

Nguyen, M.Q., Atkinson, P.M. and Lewis, H.G. (2006) 'Superresolution mapping using a Hopfield neural network with fused images', *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 44, No. 3, pp.736–749.

Nguyen, M.Q., Atkinson, P.M. and Lewis, H.G. (2011) 'Super-resolution mapping using Hopfield neural network with panchromatic imagery', *International Journal of Remote Sensing*, Vol. 32, No. 21, pp.6149–6176.

Shen, Z., Qi, J. and Wang, K. (2009) 'Modification of pixel–swapping algorithm with intialization from a sub-pixel/pixel spatial attraction model', *Photogrammetric Engineering and Remote Sensing*, Vol. 75, No. 5, pp.557–567.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S. (2001a) 'Land cover mapping at the sub-pixel scale using a Hopfield neural network', *International Journal of Applied Earth Observation and Geoinformation*, Vol. 3, No. 2, pp.184–190.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S. (2001b) 'Super–resolution mapping of urban scenes from IKONOS imagery using a Hopfield neural network', *Proceedings of the International Geoscience and Remote Sensing Symposium*, IEEE, Sydney, pp.3203–3205.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S. (2001c) 'Super-resolution target identification from remotely sensed images using a Hopfield neural network', *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 39, No. 4, pp.781–796.

Tatem, A.J., Lewis, H.G., Atkinson, P.M. and Nixon, M.S. (2003) 'Increasing the spatial resolution of agricultural land cover maps using a Hopfield neural network', *International Journal of Geographical Information Science*, Vol. 17, No. 7, pp.647–672.

Wang, P., Wang, L.G., Leung, H. and Zhang, G. (2019) 'Subpixel mapping based on hopfield neural network with more prior information', *IEEE Geoscience and Remote Sensing Letters*, Vol. 16, No. 8, pp.1284–1288.

Xu, X., Zhong, Y., Zhang, L. and Li, P. (2011) 'Remote sensing image subimage element localization based on spatial autocorrelation neural network', *Journal of Surveying and Mapping*, Vol. 40, No. 3, p.307.311.

Zhang, L., Wu, K., Zhong, Y. and Li, P. (2008) 'A new sub–pixel mapping algorithm based on a BP neural network with an observation model', *Neurocomputing*, Vol. 1, Nos. 10–12, pp.2046–2054.