

# Iterative sliced inverse regression for segmentation of ultrasound and MR images

Han-Ming Wu<sup>a</sup>, Henry Horng-Shing Lu<sup>b,\*</sup>

<sup>a</sup>*Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan, ROC*

<sup>b</sup>*Institute of Statistics, National Chiao Tung University, Hsinchu 300, Taiwan, ROC*

Received 21 July 2005; received in revised form 19 March 2007; accepted 27 April 2007

## Abstract

In this study, we propose an integrated approach based on iterative sliced inverse regression (ISIR) for the segmentation of ultrasound and magnetic resonance (MR) images. The approach integrates two stages. The first is the unsupervised clustering which combines multidimensional scaling (MDS) with K-Means. The dimension reduction based on MDS is employed to obtain fewer representative variates as input variables for K-Means. This step intends to generate the initial group labels of the training data for the second stage of supervised segmentation. We then combine the SIR with the nearest mean classifier (NMC) or the support vector machine (SVM) to iteratively update the group labels for supervised segmentation. The method of SIR is introduced by Li [Sliced inverse regression for dimension reduction. *J. Am. Stat. Assoc.* 86 (1991) 316–342] to explore the effective dimension reduction (*e.d.r.*) directions from the training data embedded in high-dimensional space. The test data are then projected onto these directions and the classifiers are further applied to classify the test data. The integrated approach based on ISIR is evaluated on simulated and clinical images, which include ultrasound and MR images. The evaluation results indicate that this approach provides an improvement of image segmentation over the methods to be compared without dimension reduction.

© 2007 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.

**Keywords:** Unsupervised clustering; K-Means; Dimension reduction; Multidimensional scaling; Sliced inverse regression; Nearest mean classifier; Support vector machines

## 1. Introduction

Segmentation is a preliminary step for further analysis of images. It is used to partition the images into regions that share similar characters. Segmentation of images with soft tissues and tumors is typically encountered in medical images like the ultrasound and magnetic resonance (MR) images. Image segmentation can be performed in two different ways, namely through unsupervised and supervised segmentation. In the case of the latter, it is required that the group labels of the training data be known a priori, while the former does not. Numerous image segmentation techniques are available in the literature. A review can be found in Clarke et al. [1] and Pal and Pal [2].

We have proposed improved snake models for ultrasound image segmentation [3–6]. However, the shape-based segmentation methods required prior information about the desired objects. Supervised techniques including neural networks [7] have been applied to segment MR images of the brain. Due to the large number of neurons used, a high computational cost is inevitable. In Ref. [8], the principal component analysis (PCA) is used to compare with other transformation methods on the human brain MR image for image segmentation. PCA is widely used in segmentation due to its advantage of easy implementation. We planned in the present work to explore the application of dimension reduction techniques beyond the PCA for image segmentation.

In this study, we propose an integrated approach based on iterative sliced inverse regression (ISIR) to segment images from the perspective of statistical classification for image pixels in the presence of spatial structure. The approach includes two stages for unsupervised clustering and supervised segmentation

\* Corresponding author. Tel.: +886 3 573 1870; fax: +886 3 572 8745.

E-mail addresses: [hmwu@stat.sinica.edu.tw](mailto:hmwu@stat.sinica.edu.tw) (H.-M. Wu),  
[hslu@stat.nctu.edu.tw](mailto:hslu@stat.nctu.edu.tw) (H.H.-S. Lu).

with dimension reduction techniques. The image data with spatial structure in high dimension is typically encountered when performing segmentation. Hence, the dimension reduction is used to overcome the curse of the dimensionality and to reduce the computational costs. The dimension reduction technique employed in our unsupervised clustering stage is multidimensional scaling (MDS) [9]. The resulting vectors obtained from training data are called MDS-variates. These variates are used as input variables for K-Means [10]. For K-Means, the number of clusters has to be specified a priori. This step is designed to obtain the initial group labels of the training data. The resulting group labels were further used to perform a supervised segmentation algorithms in the second stage.

The supervised segmentation will be conducted by a simple nearest mean classifier (NMC) or support vector machine (SVM) to iteratively update the group labels based on the sliced inverse regression (SIR) [11]. SIR finds the effective dimension-reduction (*e.d.r.*) directions of the data that provides a simple and fast algorithm to visualize the data. This method has been extended and used in various applications [12–14]. Especially, the directions obtained by SIR exhibit the largest difference among group means relative to the within-group variance, which are useful for pixel classification purposes. The method of NMC is simple which uses the label of the nearest group mean to perform classification. Recently, the support vector machines (SVM) received a lot of attention for its good performance [15,16]. It was originally designed for binary classification problem. We employed two extensions of the SVMs for multiclass data, namely the LIBSVM [17] and the SVMTorch [18] for comparison.

We evaluated the performance of the K-Means with or without dimension reduction by performing the unsupervised segmentation on the simulated and texture images in three different feature domains of images: space, frequency and space–frequency domains. Moreover, we compared the performance of the supervised segmentation by classifiers with or without dimension reduction by SIR. The classifiers consist of NMC or SVM. An iterative process is proposed to offer the advantages of eliminating the need for user intervention in order to carry out image segmentation. The flow chart of unsupervised and supervised segmentation in the proposed iterative process is shown in Fig. 1.

The paper is organized as follows. The feature extraction of images is given in Section 2. Section 3 describes the MDS as well as K-Means for unsupervised clustering procedures. Section 4 presents the proposed segmentation procedure which includes NMC or SVM in ISIR. The segmentation results on the simulated, texture and clinical medical images are reported in Sections 5 and 6. The conclusion and discussion are addressed in Section 7.

## 2. Feature extraction of images

*Space domain: Local blocks.* A rectangular lattice for a digital image in 2D of size  $N \times M$  is denoted by  $\mathcal{S} = \{(x, y) | 1 \leq x \leq N, 1 \leq y \leq M; x, y \in \mathbf{Z}\}$ , and denote the grey level of an image at some pixel  $i$  by  $f(i)$ ,  $i \in \mathcal{S}$ . The

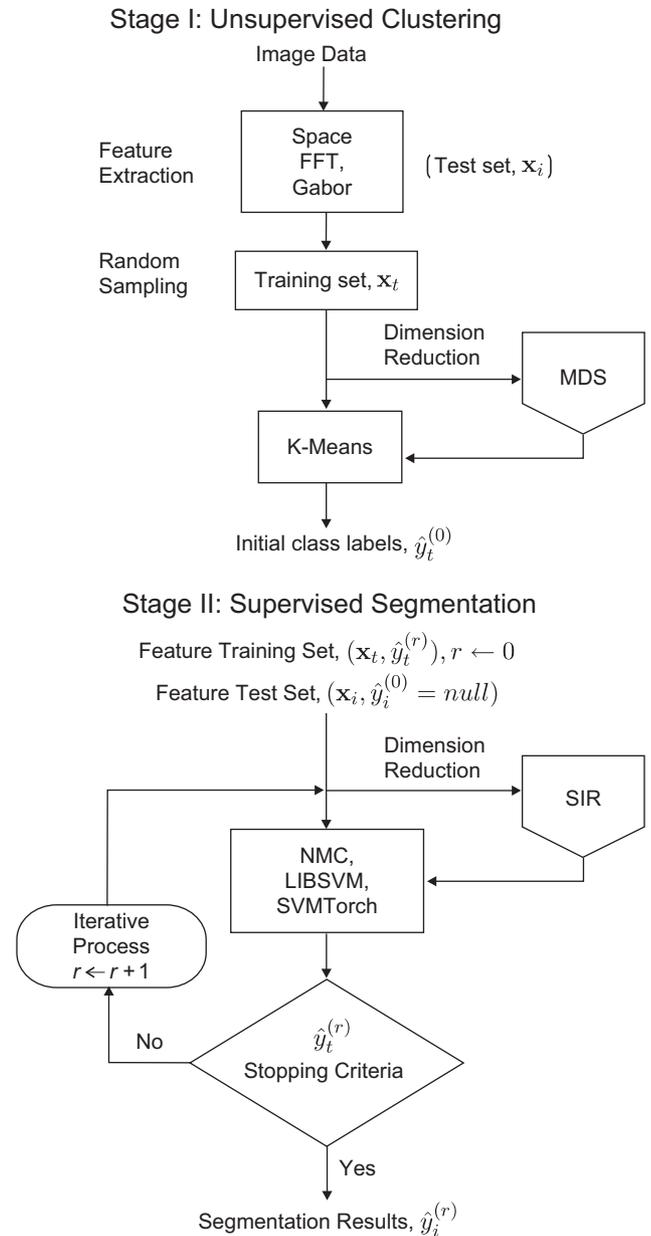


Fig. 1. The flow chart of the proposed approach for the two-stage image segmentation is illustrated.

spatial characteristic of a pixel in a image is described by its neighboring pixels. So, a local block with size  $b \times b$  can be formed as a feature vector  $\mathbf{x}_i$  of the central pixel  $i \in \mathcal{S}$ ,  $i = 1, \dots, n$  and  $n = (M - b + 1)(N - b + 1)$ . The dimension of the feature vector in the space domain is then  $p = b^2$ . These vectors capture the local spatial features of an image and are suitable for simple and regular textures. For supervised segmentation, an image with  $n$  feature vectors is a testing data, we randomly select  $m < n$  feature vectors as training data.

*Frequency domain: Fourier transform of local blocks.* If the feature of an image is periodical over space, then the feature of a local block in the space domain can be transformed to the frequency domain by the Fourier transform. This transform will highlight the periodical pattern [19]. This can be performed by

fast Fourier transform if the block size is of the power of 2. For a block with size  $b \times b$ , the two dimensional discrete Fourier transform can be expressed as

$$F(u, v) = \frac{1}{b^2} \sum_{x=0}^{b-1} \sum_{y=0}^{b-1} f(x, y) \exp \left[ -i2\pi \left( \frac{ux}{b} + \frac{vy}{b} \right) \right],$$

where  $i = \sqrt{-1}$ ,  $u, v = 0, \dots, b-1$ . Because the image intensity is real valued, the Fourier transform is symmetrical about the center. By this symmetry, almost a half of FFT calculation is redundant. Therefore, the feature vector  $\mathbf{x}_i$  in the frequency domain consists of  $|F(u, v)|$  with dimension  $p = b^2/2 + 2$  if  $b$  is power of 2. In the present study, the Fourier terms are complex and the real amplitudes are used in calculation of  $|F(u, v)|$  for the features in the frequency domain.

*Space–frequency domain: Gabor filter banks of local blocks.* Human vision has demonstrated its superior capacity in detecting boundaries of desired objects. Numerous researches on texture classification and segmentation have been carried out based on early vision models [3–5]. The main concept is to conduct segmentation on the so-called *neuroimages*. The neuroimages are constructed by convolving the observed image with a bank of specific frequencies and orientation bands, such as a bank of Gabor functions. The general form of a Gabor function is given by

$$g(x, y) = \exp\{-(x - x_0)^2 a_1^2 + (y - y_0)^2 a_2^2\} \pi \\ \times \exp\{-2\pi i[u_0(x - x_0) + v_0(y - y_0)]\},$$

and its Fourier transform is

$$G(u, v) = \frac{1}{a_1 a_2} \exp \left\{ -\frac{1}{\pi} \left[ \frac{(u - u_0)^2}{a_1^2} + \frac{(v - v_0)^2}{a_2^2} \right] \right\} \\ \times \exp\{-2\pi i[x_0(u - u_0) + y_0(v - v_0)]\},$$

where  $(a_1, a_2)$  are parameters for scaling the two axis of the Gaussian envelop,  $(x_0, y_0)$  are location parameters of the peak of the Gaussian envelop, and  $(u_0, v_0)$  are spatial frequencies of the sinusoidal in Cartesian coordinates. Each local block is convolved with a bank of Gabor filters with different frequencies and orientations. We employed the so-called *G-vector* as the feature vector [6] at pixel  $i = (x, y)$ , which is computed by

$$\mathbf{x}_i = \{g_{pk}(x, y), g_{nk}(x, y); k = 1, \dots, r\},$$

where  $g_{pk}(x, y)$  and  $g_{nk}(x, y)$  are the summations of the positive and negative values for the neuroimages that is the convoluted image of the  $k$ th Gabor filter. Thus, the dimension of the feature vector is  $p = 2r$  in the space–frequency domain. For instance, we can consider a bank of  $r = 4 \times 6 = 24$  Gabor filters that are designed with four scale of center frequencies  $\sqrt{2}/2, \sqrt{2}, 2\sqrt{2}$  and  $b\sqrt{2}/4 = 4\sqrt{2}$  when  $b = 16$ , as well as six orientation of angles of  $0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ$  and  $150^\circ$ .

### 3. Unsupervised clustering by K-Means on MDS-variates

Let  $X = \{\mathbf{x}_i\}_1^n$  be a data set with  $n$  objects, where  $\mathbf{x}_i$  is a feature vector in a  $p$ -dimensional metric space. Cluster analysis results

in non-overlapped  $K$  groups of a data set in which observations preserve the coherent patterns. K-Means clustering [10] is one of most well known clustering methods. Class labels  $\hat{y}_i$ 's are determined for the feature vectors  $\mathbf{x}_i$ 's by computing the centroid  $\bar{\mathbf{x}}_k$  for each group and assigning each observation to the group with the closest centroid. That is,  $\hat{y}_i = \arg \min_k \|\mathbf{x}_i - \bar{\mathbf{x}}_k\|$ . K-Means is an iterative procedure because once the new  $\bar{\mathbf{x}}_k$  is obtained,  $\hat{y}_i$  changes and need to be recalculated which in turn affect  $\bar{\mathbf{x}}_k$ . These steps are repeated until  $\bar{\mathbf{x}}_k$  stabilized.

The final clustering results highly depend on the initial  $\bar{\mathbf{x}}_k$ . In this study, the initial group centroids are obtained by first ordering the distances of data points to the sample mean. Then, the  $\{1 + n(k-1)/K\}$ th point is selected as the initial centroid for group  $k$ , where  $n$  is the total number of data points. The number of groups  $K$  has to be known in the algorithm. In the present study,  $K$  is specified according to the scientific problem of interest. For example, the tumor and non-tumor tissues in ultrasound images form two groups. The soft tissues and other non-soft tissues in MRI images form two distinct groups.

It is common to first extract features from  $X$  suitable for segmentation before applying the clustering or classification algorithm at hand. Multidimensional scaling proposed by Torgerson [20] is one of such useful feature extractors. MDS presents a configuration of  $n$  objects in a low-dimensional Euclidean space based on the information about the distance between these objects. Classical MDS corresponds to linear transformation of the input variables. The resulting combinations are called MDS-variates. MDS bears much similarity to two other unsupervised linear projection methods, namely, PCA and factor analysis. We combine the MDS with K-Means to perform the proposed clustering procedure. Applying the MDS to the training data will produce the MDS-variates in the lower dimensional subspace. The training data are image features  $\mathbf{x}_i$ 's in the space, frequency and space–frequency domains. K-Means can be used to cluster these variates and assign a group label to each observation in the training data. The resulting clusters are the initial labels for further supervised segmentation. We denote the proposed unsupervised clustering approach as MDS + K-Means.

### 4. Supervised segmentation

In the supervised segmentation, it is assumed that a training data with known group labels are available. In the present study, a training data of size  $m$  is constructed by randomly sampling from  $n$  testing feature vectors, where the training group labels  $\{\hat{y}_t\}_1^m$  for the training feature vectors  $\{\mathbf{x}_t\}_1^m$  are obtained in the preceding step. The training data is used to perform the SIR algorithm for finding the projection directions. These projected data are used to construct the classifiers and perform the segmentation.

#### 4.1. Iterative sliced inverse regression (ISIR)

Li [11] presented the method of sliced inverse regression (SIR) as a prototypical framework for dimension reduction.

The following regression model is considered:

$$y = f(\beta_1^t \mathbf{x}, \dots, \beta_B^t \mathbf{x}, \varepsilon), \quad (1)$$

where  $y$  is a univariate variable,  $\mathbf{x}$  is a random vector with dimension  $p \times 1$ ,  $p \geq B$ ,  $\beta$ 's are vectors with dimension  $p \times 1$ ,  $\varepsilon$  is a random variable independent of  $\mathbf{x}$  and  $f$  is an arbitrary function. The  $\beta$ 's are referred to as *e.d.r.* directions or SIR projection directions to be estimated. SIR is a method for estimating the *e.d.r.* directions based on  $y$  and  $\mathbf{x}$ . Under regular conditions in Ref. [11], it is shown that the centered inverse regression curve  $E[\mathbf{x}|y] - E[\mathbf{x}]$  is contained in the linear subspace spanned by  $\beta_j^t \Sigma_{xx}$  ( $j = 1, \dots, B$ ), where  $\Sigma_{xx}$  denotes the covariance matrix of  $\mathbf{x}$ . Based on these facts, the estimated  $\beta$ 's can be obtained by the procedures of standardizing  $\mathbf{x}$ , partitioning slices (or groups) according to the value of  $y$ , calculating the slice means of  $\mathbf{x}$ , and performing the PCA of the slice means with weights as sample proportion within each slice.

The group labels obtained by the proposed clustering methods are the initial partitions. Generally, the clustering results are not precise enough to obtain better segmentation. Further improvements by iteration of supervised classification could be applied to improve the segmentation results. We propose an iterative SIR (ISIR) algorithm for estimating the *e.d.r.* directions and image segmentation based on the training data  $\{\hat{y}_t, \mathbf{x}_t\}_1^m$ . This approach can be considered to be analogous to the K-Means in spirit which iteratively updates the group labels to construct the SIR projection directions. We could use these new projection directions for further segmentation iteratively. The algorithm is described as follows.

*Algorithm of ISIR:*

1. Compute sample mean  $\bar{\mathbf{x}}$  and sample covariance matrix  $\hat{\Sigma}_{xx}$  of  $\mathbf{x}_t$ 's.
2. Divided the data set into  $K$  slices according to the categories of  $\hat{y}_t$ . Denoted these slices by  $I_1, \dots, I_K$ . Let the proportion of all observed  $\hat{y}_t$ 's that falls in  $k$ th slice be  $p_k$ , i.e.,

$$p_k = \frac{\sum_{t=1}^m \delta_k(\hat{y}_t)}{m} = \frac{m_k}{m}, \quad \delta_k(\hat{y}_t) = 1 \quad \text{if}$$

$$\hat{y}_t \in I_k, \quad \delta_k(\hat{y}_t) = 0 \quad \text{otherwise.}$$

3. Within each slice, compute the sample mean  $\bar{\mathbf{x}}_k$  of  $\mathbf{x}_t$ 's,  $k = 1, \dots, K$ , and the weighted covariance matrix,  $\hat{\Sigma}_W = \sum_{k=1}^K p_k (\bar{\mathbf{x}}_k - \bar{\mathbf{x}})(\bar{\mathbf{x}}_k - \bar{\mathbf{x}})^t$ .
4. Find the eigenvalues and eigenvectors for  $\hat{\Sigma}_W$  with respect to  $\hat{\Sigma}_{xx}$  by solving

$$\hat{\Sigma}_W \hat{\beta}_j = \lambda_j \hat{\Sigma}_{xx} \hat{\beta}_j,$$

where  $j = 1, \dots, p$  and  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ . Then, the leading  $B$  eigenvectors  $\hat{\beta}_j$ 's are used as the projection directions.

5. A projected training data and testing data is obtained by  $\mathbf{z}_t = (\hat{\beta}_1^t \mathbf{x}_t, \dots, \hat{\beta}_B^t \mathbf{x}_t)$  and  $\mathbf{z}_i = (\hat{\beta}_1^i \mathbf{x}_i, \dots, \hat{\beta}_B^i \mathbf{x}_i)$ . We call  $\mathbf{z}_i$ 's the SIR-variates. A classifier is constructed using  $\{\hat{y}_t, \mathbf{z}_t\}_1^m$ . Denote the predicted group labels for training data and

testing data by  $\hat{y}_t^{(r)}$ , and  $\hat{y}_i^{(r)}$  for the current iteration  $r$ . If  $r = 1$ , this refers to the classical SIR approach.

6. *Iterative process:* repeat step 2–5 unless one of the following criteria holds:

$$\left\{ \left( \sum_{t=1}^m \mathbf{I}[\hat{y}_t^{(r)} \neq \hat{y}_t] \right) - \left( \sum_{t=1}^m \mathbf{I}[\hat{y}_t^{(r-1)} \neq \hat{y}_t] \right) \right\} / m \leq \varepsilon (=0.001) \quad \text{or} \quad (2)$$

$$\sum_{t=1}^m \mathbf{I}[\hat{y}_t^{(r)} \neq \hat{y}_t] / m > \eta (=0.5) \quad \text{or} \quad (3)$$

$$r > R (=100), \quad (4)$$

where  $\mathbf{I}$  is an indicator function with  $\mathbf{I}[x] = 1$  if  $x$  holds, and zero otherwise.

Criterion (2) describes the error rate difference between the successive segmentations is within a tolerance  $\varepsilon$  such as 0.001, in which the result reaches a stable segmentation. Criterion (3) states that the mean difference between the initial and the  $r$ th result cannot exceed  $\eta$ , like 50%. This may happen when the algorithm is diverged and the final segmentation is far from the initial. If this criterion holds, the final segmentation is determined at the  $\tilde{r}$ th iteration with the minimum mean difference. That is,  $\tilde{r} = \arg \min_r \sum_{t=1}^m \mathbf{I}[\hat{y}_t^{(r)} \neq \hat{y}_t] / m$ . Criterion (4) limits the number of iterations within a fixed number  $R$ , like 100, to avoid an endless loop when the other two criteria fail to hold.

According to Theorem (5.1) in Ref. [11], the average of the smallest  $p - B$  eigenvalues has asymptotically a  $\chi^2$  distribution with  $(p - B)(K - B - 1)$  degrees of freedom if  $\mathbf{x}$  has a normal distribution. As a result, the number of directions  $B$  is less than or equal to  $\min\{K - 1, p\}$ . The ISIR algorithm can be applied to obtain the stable *e.d.r.* directions,  $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_B)$ , where  $B \leq \min\{K - 1, p\}$ . The responses  $y$ 's in the image are of a finite number of  $K$  categories. Those  $(K - 1)$  *e.d.r.* directions are used for segmentation of  $K$  regions because  $K - 1$  is typically smaller than or equal to  $p$ . Further discussions regarding the determination of  $K$  are reported in Refs. [11,21]. However, SIR algorithm will fail to find the *e.d.r.* directions when the data is symmetric. In this case, the principal Hessian directions (pHd) [22] can be used as an alternative method to find the *e.d.r.* directions.

Similar to PCA, SIR is a method based on the projection of input variables  $\mathbf{x}$  to the latent variables (components). However, in contrast to PCA, SIR creates the components by modeling the relationship between input  $\mathbf{x}$  and response variables  $y$  while maintaining most of the information in the input variables. SIR can be seen as a PCA-like procedure performed on the random variable  $E(\mathbf{x}|y)$  instead of on  $\mathbf{x}$ . That is, SIR looks for linear combinations of  $\mathbf{x}$  which maximize  $\text{var}(E(\mathbf{a}^t \mathbf{x}|y)) / \text{var}(\mathbf{a}^t \mathbf{x})$  instead of just  $\text{var}(\mathbf{a}^t \mathbf{x})$ . While PCA leads to an eigen-system, SIR leads to a generalized eigen-system.

The leading canonical variates in Fisher linear discriminant analysis (LDA) are the linear combinations of  $\mathbf{x}$  formed by the vector  $\mathbf{a}$ , which solves the maximization problem by  $\max_{\mathbf{a}} (\mathbf{a}^t \Sigma_B \mathbf{a} / \mathbf{a}^t \Sigma_W \mathbf{a})$ , where  $\Sigma_B$  and  $\Sigma_W$  are the between-group and within-group covariances, respectively. As described

in Ref. [23], the SIR variates are the same as the Fisher canonical variates except for possible difference in scale. That is, the LDA and SIR are equivalent in the population. Differences can arise in practice depending on the methods used to estimate  $\Sigma_B$  and  $\Sigma_W$ . More discussion on the connection between SIR and LDA was addressed in Ref. [24].

#### 4.2. Nearest-mean classifier (NMC)

Nearest mean classifier (NMC) [25] is a simple and widely used classifier. We use NMC for its simplicity and robustness in this study. The NMC calculates the mean of the training vectors for each group and perform classification based on the nearest distance of a feature vector to group means. When NMC is used with SIR, we can first project feature vectors onto these *e.d.r.* directions found by SIR and then apply NMC on projected feature vectors. The mean of the projected feature vectors for every group in the training set is the centroid. A pixel  $i$  in a test (or training) image is classified (or predicted) into the  $k$ th group if the projected feature vector of that pixel is closest to the centroid of  $k$ th group. That is, the predicted group label for a feature vector  $\mathbf{x}_i$  is given by

$$\hat{y}_i = \arg \min_k \|\beta' \mathbf{x}_i - \bar{\mathbf{z}}_k\|, \quad i = 1, \dots, n, \quad (5)$$

where  $\bar{\mathbf{z}}_k = \sum_{t \in I_k} \mathbf{z}_t / m_k$  and  $\mathbf{z}_t = \beta' \mathbf{x}_t$ .

#### 4.3. Support vector machines (SVMs)

SVMs have been researched and applied to many problems recently [26,27]. This is particularly for its theoretical foundations in computational learning theory and improved performances in applications. The fundamental concept is to map the data into high-dimensional space through a kernel function  $\phi$  and then find a hyperplane  $\mathbf{w}$  to separate the groups. It was originally designed for binary separation. To extend this method to multiclass classification, two approaches can be implemented. One is the “one-against-the-others” approach, in which the  $k$ th SVM model is constructed with all of the samples belonging to the  $k$ th group in one group, and all the other samples with another group. This is implemented in SVMTorch [18]. The  $k$ th SVM solves the following optimization problem:

$$\min_{\mathbf{w}_k, b_k, \xi_k} \frac{1}{2} \|\mathbf{w}_k\|^2 + C \sum_{t=1}^m \xi_{kt},$$

$$\mathbf{w}_k^t \phi(\mathbf{x}_t) + b_k \geq 1 - \xi_{kt} \quad \text{if } y_t = k,$$

$$\mathbf{w}_k^t \phi(\mathbf{x}_t) + b_k \leq -1 + \xi_{kt} \quad \text{if } y_t \neq k,$$

$$\xi_{kt} \geq 0, \quad t = 1, \dots, m,$$

where  $\phi$  is a function and  $C$  is the penalty parameter. The decision functions can be expressed as  $\mathbf{w}_k^t \phi(\mathbf{x}) + b_k$ ,  $k = 1, \dots, K$ . The predicted group label  $\hat{y}_i$  for a pixel  $\mathbf{x}_i$  in the test data is given by

$$\hat{y}_i = \arg \max_k \{\mathbf{w}_k^t \phi(\mathbf{x}_i) + b_k, \quad k = 1, \dots, K\}. \quad (6)$$

The second one is the “one-against-one” approach. The SVM trained model is constructed by using any of the two groups. Hence, there are a total of  $K(K-1)/2$  classifiers. It is implemented in LIBSVM [17] which solves the following binary classification problem:

$$\min_{\mathbf{w}_{kl}, b_{kl}, \xi_{kl}} \frac{1}{2} \|\mathbf{w}_{kl}\|^2 + C \sum_{t=1}^m \xi_{klt},$$

$$\mathbf{w}_{kl}^t \phi(\mathbf{x}_t) + b_{kl} \geq 1 - \xi_{klt} \quad \text{if } y_t = k,$$

$$\mathbf{w}_{kl}^t \phi(\mathbf{x}_t) + b_{kl} \leq -1 + \xi_{klt} \quad \text{if } y_t = l,$$

$$\xi_{klt} \geq 0, \quad t = 1, \dots, m,$$

where the training data are from the  $k$ th and  $l$ th groups. There will be  $K(K-1)/2$  classifiers to be constructed with the decision function:

$$\mathcal{C}_g(\mathbf{x}_i) = \begin{cases} k, & \mathbf{w}_{kl}^t \phi(\mathbf{x}_i) + b_{kl} \geq 0, \\ l & \text{otherwise.} \end{cases}$$

$$g = 1, \dots, K(K-1)/2.$$

A pixel  $\mathbf{x}_i$  in the test sample is classified into  $\hat{y}_i$  class by majority voting, that is

$$\hat{y}_i = \arg \max_k \left\{ \sum_{g=1}^{K(K-1)/2} \mathbf{I}[\mathcal{C}_g(\mathbf{x}_i) = k], \quad k = 1, \dots, K \right\},$$

(7)

where  $\mathbf{I}$  is an indicator function. If two groups have identical votes, LIBSVM selects one with the smaller index. These two types of support vector machines are used to demonstrate the proposed segmentation methods. In the present study, all parameters used the default settings. Using different parameters could lead to different results which can be evaluated in the follow-up studies.

In the following, the segmentation approaches based on input feature vectors  $\mathbf{x}_i$  without dimension reduction are denoted as NMC, LIBSVM, and SVMTorch. The extended segmentation approaches with SIR-variates  $\mathbf{z}_i$  are namely SIR+NMC, SIR+LIBSVM and SIR+SVMTorch.

## 5. Simulation studies

In this section, we first perform the classical SIR+NMC, SIR+LIBSVM and SIR+SVMTorch methods for segmentation of the simulated images and the texture images. The aim was to illustrate the ability and usefulness of the classical SIR methods in the image segmentation problem. We then performed the iterative process on the image in Fig. 4(c) to contrast the relative effectiveness of the iterative SIR methods to the classical one.

### 5.1. Simulated images

For simplicity, we consider the case in which we constructed a synthetic image with size  $64 \times 128$  as shown in Fig. 2(a).

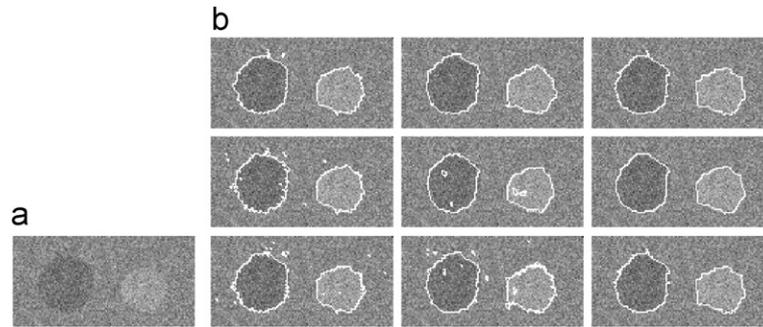


Fig. 2. (a) The simulated image is used for evaluation of the proposed approach which consists of two circle objects (Cb, Cs) and background (Bg). (b) The segmentation results by the SIR + NMC, SIR + LIBSVM and SIR + SVM Torch from top to bottom are demonstrated based on the clustering results of MDS + K-Means with a block size of  $8 \times 8$ . The images from left to right are performed in the space, frequency and space–frequency domains, respectively.

There are three groups, namely two circle objects (denoted by Cb, Cs) and a background (denoted by Bg) with grey levels 80, 120 and 100, respectively. The added noises are randomly distributed from a normal distribution with mean zero and the standard deviation of 20. Hence, the signal-to-noise ratio (SNR) is equal to one. One thousand training samples are randomly chosen for unsupervised clustering as one run. Table 1 shows the average clustering error rates of 50 runs with various block size in the space domain. One can see that the errors were reduced while the block size is  $9 \times 9$  in this case. Since the algorithm of FFT is fast for the block size in the power of 2, we consider the block sizes of  $4 \times 4$  or  $8 \times 8$  in this case for comparing the effect of the block size.

Table 1  
The average clustering error rates with various block sizes in the space domain

Block size	K-Means	MDS + K-Means
$4 \times 4$	0.0633	0.0625
$5 \times 5$	0.0447	0.0438
$8 \times 8$	0.0432	0.0383
$9 \times 9$	0.0372	0.0364
$16 \times 16$	0.0410	0.0421

Table 2 presents the average error rates of 50 runs by the methods of K-Means and MDS + K-Means with a block size of  $4 \times 4$  and  $8 \times 8$ . We can observe that the average clustering error rates with a block size of  $8 \times 8$  are lower than those with a block size of  $4 \times 4$  in this case whenever the MDS was employed. The MDS-variables can reduce the error rate of K-Means, particularly with a block size of  $8 \times 8$ . The performance in the space–frequency domain is better than those in the other two domains. The best performance is obtained by the method of MDS + K-Means in the space–frequency domain with a block size of  $8 \times 8$ . This gave a good training set with initial group labels for further supervised segmentation.

To benchmark the performance of supervised segmentation in this simulation study, true group labels in the training set are also used as initial group labels to estimate segmentation errors for different classifiers when the true group labels are available *a priori*. Fig. 3 presents the average segmentation error rates of 50 runs based on the segmentation results of different classifiers using the true group labels in the training set or the initial group labels generated by the methods of K-Means and MDS + K-Means. We summarize comparative results for this

Table 2  
The average error rates for the training data of the simulated image by the clustering methods are reported with a block size of  $4 \times 4$  and  $8 \times 8$  in the space, frequency, and space–frequency domains, where  $p$  represents the feature dimension

Clustering methods	Block size	Features	$p$	Cs	Cb	Bg	Total
K-Means	$4 \times 4$	Space	16	0.0732	0.0574	0.0594	0.0633
		FFT	10	0.0712	0.0553	0.0593	0.0619
		Gabor	48	0.0331	0.0517	0.0509	0.0452
	$8 \times 8$	Space	64	0.0163	0.0685	0.0448	0.0432
		FFT	34	0.0169	0.0671	0.0411	0.0417
		Gabor	48	0.0223	0.0568	0.0355	0.0382
MDS + K-Means	$4 \times 4$	Space	1	0.0747	0.0561	0.0568	0.0625
		FFT	1	0.0706	0.0545	0.0589	0.0613
		Gabor	1	0.0326	0.0459	0.0511	0.0432
	$8 \times 8$	Space	1	0.0169	0.0663	0.0316	0.0383
		FFT	1	0.0172	0.0673	0.0282	0.0376
		Gabor	1	0.0222	0.0538	0.0242	0.0334

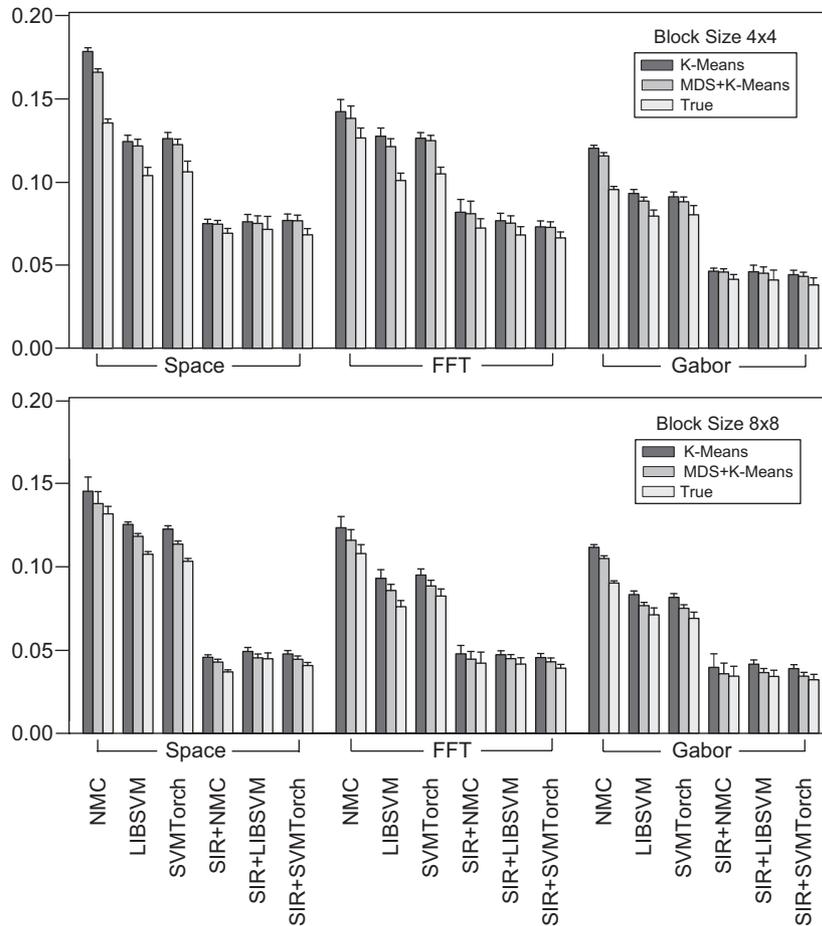


Fig. 3. The average error rates with one standard deviation bars of 50 runs in segmentation are plotted by the segmentation results of different classifiers using the true group labels in the training set or the initial group labels generated by the methods of K-Means and MDS + K-Means.

case as follows:

1. Apparently, the SIR versions of the segmentation methods reduce the error rates significantly.
2. By employing the Gabor features in the space–frequency domain, the performances for the segmentation methods have the lower error rates than those with the other two features in the space or frequency domains.
3. The lowest total errors of 3.82% and 3.22% appear in SIR + SVMTorch with Gabor features based on the segmentation results of MDS + K-Means with the block size of  $4 \times 4$  and  $8 \times 8$ , respectively.
4. The SIR can improve the performance of the segmentation methods to the extent that the simple algorithm of NMC can perform as well as the complicated algorithms of SVMs do.
5. The error rates of segmentation methods with SIR by the block size of  $8 \times 8$  are lower than those by the block size of  $4 \times 4$ .
6. The classifier of SVMTorch has the similar performance to that of LIBSVM.

Thus, the segmentation methods with SIR have the lowest errors in all cases. These preliminary studies on the simulated image have shown that the simple algorithm such as the SIR + NMC

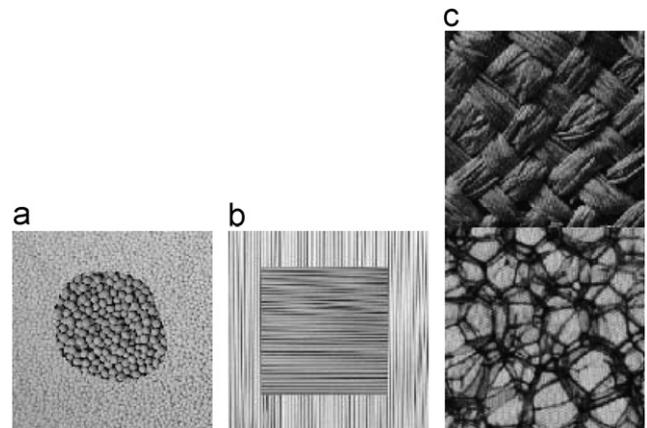


Fig. 4. The test texture images (a), (b) and (c) are displayed.

is compatible with the complicated algorithms such as the SIR + LIBSVM and the SIR + SVMTorch. The results of segmented images by the SIR + NMC, SIR + LIBSVM and SIR + SVMTorch based on the clustering results of MDS + K-Means with a block size of  $8 \times 8$  are shown in Fig. 2(b). The boundaries of the segmentation are superimposed on the original test image for visualization.

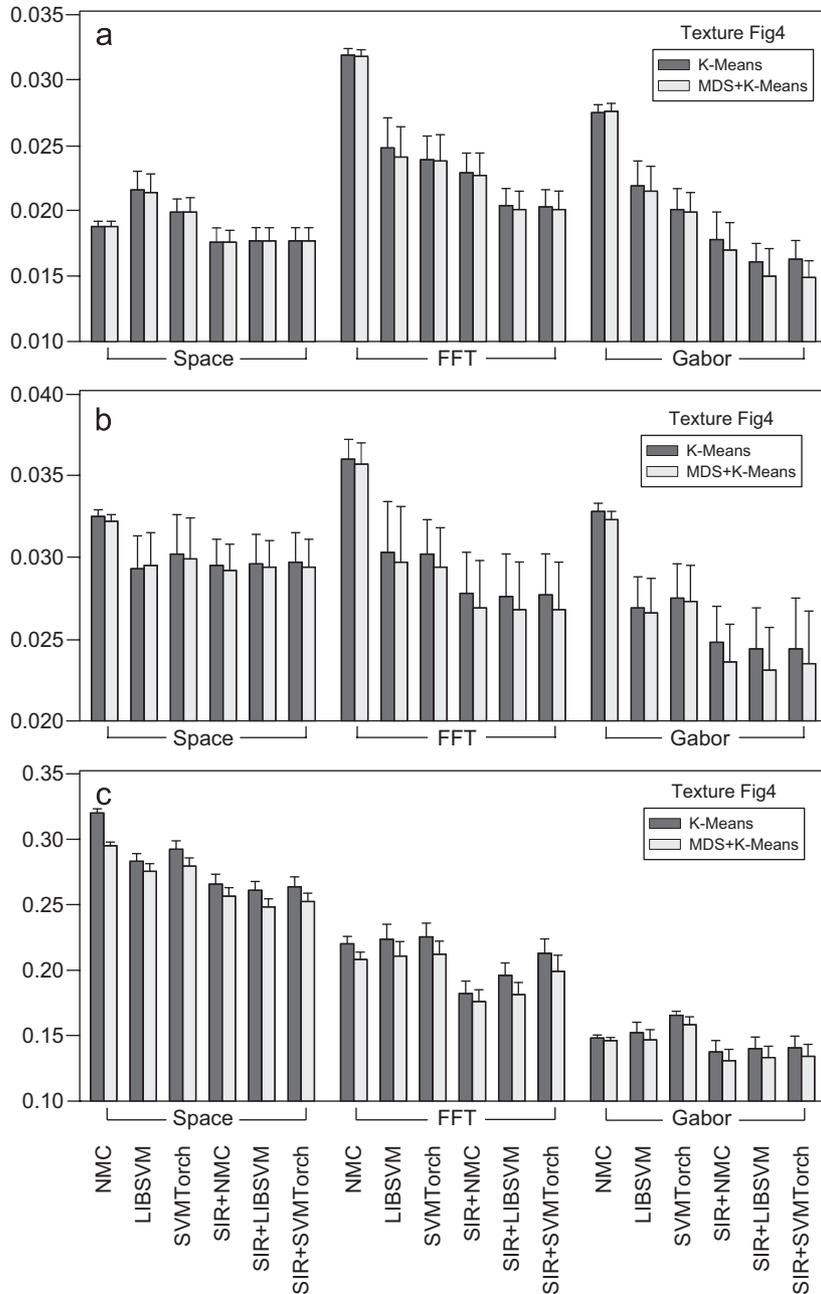


Fig. 5. The average error rates with one standard deviation bars of 50 runs in segmentation are plotted for the texture images Fig. 4(a), (b) and (c) based on the initial group labels generated by K-Means or MDS + K-Means.

5.2. Texture images

We demonstrated the proposed segmentation methods on three texture images displayed in Fig. 4. There are two groups in every image, which are denoted by Tf and Tb. One-thousand samples were randomly selected from one image to form the training data. The block sizes of  $8 \times 8$ ,  $8 \times 8$ , and  $16 \times 16$  are chosen for texture images in Fig. 4(a), (b) and (c), respectively. Fig. 5 shows the average error rates of 50 runs in segmentation based on the results of K-Means and MDS + K-Means. The results demonstrated that those segmentation methods with SIR improved the performance significantly. Fig. 6 show the

segmentation results by the SIR + NMC, SIR + LIBSVM and SIR + SVMTorch based on the clustering results of the MDS + K-Means with a block size of  $8 \times 8$  in the space–frequency domain. The left panel of Fig. 7 shows the SIR + NMC segmentation for the texture image Fig. 4(c) based on the MDS + K-Means in the frequency domain. Apparently, the texture image Fig. 4(c) is difficult to segment because of the similarity of textures in two regions. We then applied the iterative SIR + NMC on it in the frequency domain. The segmentation error rates could be reduced from 0.2003 to 0.0272 after 13 iterations as shown in the right panel of Fig. 7. The final segmented image can locate the target boundary correctly. From

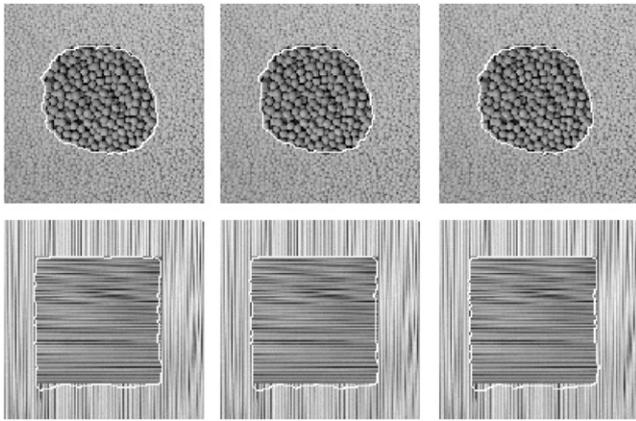


Fig. 6. The segmentation results by the SIR + NMC, SIR + LIBSVM and SIR + SVMTorch (from left to right) are displayed based on the clustering results of the MDS + K-Means for the texture images Figs. 4(a) and 4(b) with a block size of  $8 \times 8$  in the space–frequency domain.

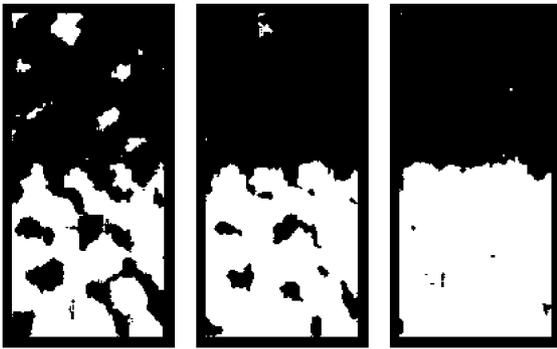


Fig. 7. The ISIR + NMC segmentation for the texture image Fig. 4(c) is demonstrated for the initial (left), five (middle) and 13 (right) iterations based on the MDS + K-Means in the frequency domain.

these simulations, the differences of error rates among successive iterations are typically decreasing for the ISIR + NMC method. However, the iterative process may not change much the segmented results when using the ISIR + LIBSVM and ISIR + SVMTorch methods. One could try different kernels and parameters in the training of SVM models to obtain improved results. For example, the segmentation error rates for the texture image in Fig. 4(a) could be reduced from 0.0201 to 0.0161 when the SIR + LIBSVM was applied with gaussian kernel and gamma 0.01 in the frequency domain based on the clustering results of MDS + K-Means.

## 6. Empirical results

For the following empirical studies of ultrasound and MR images, the iterative SIR algorithm was employed.

### 6.1. Clinical ultrasound images

The ultrasound images are widely used in the clinical diagnosis for their portability, non-invasiveness and other

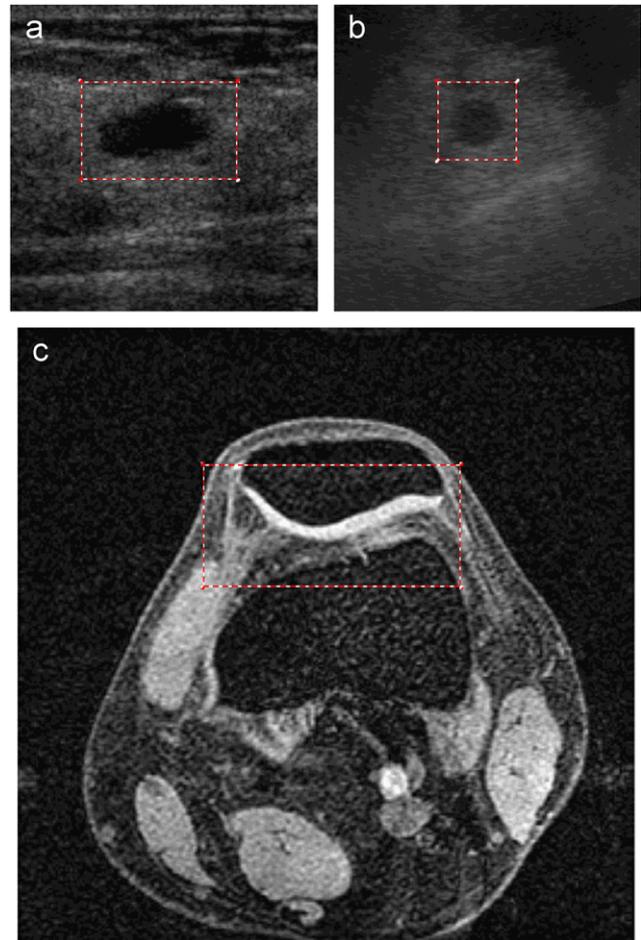


Fig. 8. The ultrasound image of (a) a breast and (b) a liver are displayed, the tumor is the object-of-interest. (b) The MR image of a knee is displayed, the cartilaginous tissues (the brighter part) is the object-of-interest. The rectangular box is ROI.

advantages. However, segmentation of ultrasound images is difficult for its low image quality, intrinsic noisy and textural properties. We use breast and liver ultrasound images to test the proposed approach. Fig. 8(a) and (b) show the ultrasound images of the breast and the liver. They were collected from the National Taiwan University Hospital. The rectangular box superimposed on the images are the ROIs and the objects of interest are the tumors. The training data consist of 1000 random samples from ROI with a block size of  $8 \times 8$ . There are two groups, namely the tumor and the non-tumor. From previous studies on simulated and texture images, we first employed the MDS + K-Means to get the initial group labels for the training data. Then we applied the ISIR + NMC, ISIR + LIBSVM and ISIR + SVMTorch methods in the space–frequency domains. The numbers of iterations were 4, 3, and 3 for Fig. 8(a) and 3, 6, and 4 for Fig. 8(b), respectively. Since there are two groups, only one SIR-variate exists. The segmented results are superimposed on the original images for comparison purposes in Fig. 9. One spot was left unfilled in the right bottom corner of the segmented breast images. This may be due to the effect of speckle noise. Overall, the boundaries of the

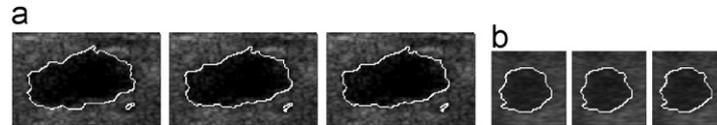


Fig. 9. The final segmentation results from left to right by the ISIR + NMC, ISIR + LIBSVM, and ISIR + SVM Torch on (a) the breast and (b) the liver ultrasound image are drawn with a block size of  $8 \times 8$  in the space–frequency domains.

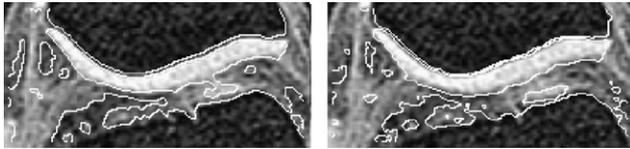


Fig. 10. The initial (left) and the final (right, after 22 iterations) segmentation results of the ISIR + NMC are demonstrated based on the MDS + K-Means with a block size of  $8 \times 8$  in the space–frequency domain.

desired targets can be successfully located by the proposed approaches.

## 6.2. Magnetic resonance images

The merits of MR images are in its high-quality contrasts of soft tissues, non-invasiveness etc. Nevertheless, the segmentation remains difficult due to the shape and intensity variation within the same structure group. Fig. 8(c) shows the MR image of a cross-section of a normal knee with an image size of  $512 \times 512$ . The rectangular box is the region of interest that contains the cartilaginous tissue (the brighter part). The objective is to segment the cartilaginous from the others. Even though the knee MR image provides a high contrast of the tissue against the other tissues, the boundary between the upper and the lower cartilaginous tissues is not clear-cut. This makes the task challenging. For diagnostic purposes, results are useful for the detection of abnormalities or damages on the knee. The classical SIR + NMC were applied on the image in the space–frequency domains and the segmentation results are shown in Fig. 10(a). It can be observed that the cartilaginous tissue could not be segmented completely with the connections remaining intact. We used the iterative SIR + NMC to get the final segmentation after 22 iterations as shown in Fig. 10(b). This example demonstrated that by applying the SIR + NMC iteratively to update the group labels could lead to the desired results.

The MR images contain a volume of images of the knee. A total of 20 2D MR images are processed individually. By using these segmented images, we can construct a 2D projection view of a knee which shows the thinness of the upper cartilaginous tissues as shown in Fig. 11. This projection view can detect the patterns of abnormalities of a knee in order to plan the operations and visualize the other anomalies. The iterative SIR approach could adjust the projection directions and reach stable segmented results to improve the diagnostic potential of medical imaging.

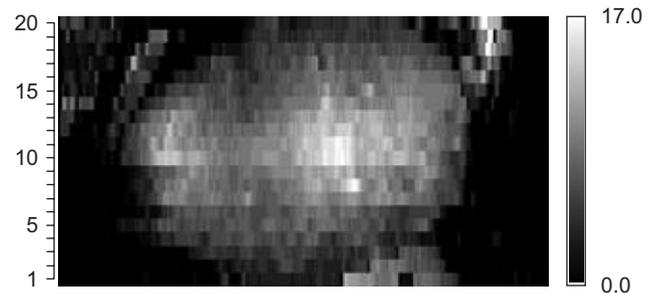


Fig. 11. The 2D projection view of the thinness of the upper cartilaginous tissues in a normal kneecap is displayed.

## 7. Conclusion and discussion

An integrated approach for images segmentation based on the dimension reduction (MDS and SIR) techniques was introduced in the present study. The K-Means based on MDS-variates is used for unsupervised clustering for the first stage. The resulting group labels are the initial estimates for the supervised segmentation in the second stage. The MDS in this step helps to reduce the errors of initial labels. The SIR with NMC or SVMs are used for the approach of ISIR in supervised segmentation. The proposed approach was evaluated on simulated, texture and clinical images in three feature domains. The findings indicate that the proposed approach significantly improved the performance of the classifiers. In addition, Gabor features could lead to successful segmentation results for complex texture images.

Unlike PCA, the gain by using SIR as a dimension reduction is that SIR adopts the group information for estimating the projection directions. These directions exhibit the largest difference in the group means relative to the within-group variance and are superior for classification problems [28]. We have proposed the ISIR method to update the group labels in order to improve the initial clusters. This suggests that an improvement of segmentation can be achieved through the ISIR method for finding the stable group memberships that behaves like the K-Means in the *e.d.r.* projection space. Hence, this method provides a simple and effective approach for image segmentation.

In this report, we use  $(K - 1)$  dimension for MDS-variates and SIR-variates to perform the segmentation algorithms of  $K$  regions. Determining the sufficient number of variates that are involved in the MDS and SIR will be explored in future studies. The selection of the block size for each image is crucial to the segmented results. Typically, the smaller the block size is, the noisier the segment becomes. However, if the block size is too

large, then the spatial structure can change. The appropriate block size can be decided by the classification error in the training set, where the training set can be obtained by hand-labeling of a medical expert. It is also noted that the different parameter settings used in SVMs could lead to different segmented results. For practical purposes, one could train the SVM models by the training data to learn the parameters. Once the SVM model is learned, we could use it to predict the test data. The issue of the stability and robustness of the proposed approach can be investigated with more images in future studies. In the future, we plan to investigate the integration of the methods proposed in the present study and the dynamic SIR model proposed in our previous study [29] to segment dynamic images.

### Acknowledgments

The authors thank Dr. Chung-Ming Chen at National Taiwan University Hospital, Dr. Hui-Cheng Cheng at Veterans General Hospital, Taipei, Taiwan, for providing and explaining the clinical ultrasound and MR images analyzed herein. The authors also thank the associate editor and referees for their valuable suggestions and reviews. This research was supported by the grants from National Science Council at Taiwan, ROC.

### References

- [1] L.P. Clarke, R.P. Velthuisen, M.A. Camacho, J.J. Heine, M. Vaidyanathan, L.O. Hall, R.W. Thatcher, M.L. Silbiger, MRI segmentation: methods applications, *Magn. Reson. Imaging* 13 (1995) 343–368.
- [2] N.R. Pal, S.K. Pal, A review on image segmentation techniques, *Pattern Recognition* 26 (1993) 1277–1294.
- [3] C.M. Chen, H.H.-S. Lu, Y.C. Lin, An early vision based snake model for ultrasound image segmentation, *Ultrasound Med. Biol.* 26 (1999) 273–285.
- [4] C.M. Chen, H.H.-S. Lu, An adaptive snake model for ultrasound image segmentation: modified trimmed mean filter, ramp integration and adaptive weighting parameters, *Ultrasonic Imaging* 22 (2001) 214–236.
- [5] C.M. Chen, H.H.-S. Lu, K.C. Han, A textural approach based on Gabor functions for texture edge detection in ultrasound images, *Ultrasound Med. Biol.* 27 (2001) 513–534.
- [6] C.M. Chen, H.H.-S. Lu, Y.S. Huang, Cell-based dual snake model: a new approach to extracting highly winding boundaries in the ultrasound images, *Ultrasound Med. Biol.* 28 (2002) 1061–1073.
- [7] S.C. Amartur, D. Piraino, Y. Takefuji, Optimization neural networks for the segmentation of magnetic resonance images, *IEEE Trans. Med. Imaging* 11 (1992) 215–222.
- [8] H. Soltanian-Zadeh, J.P. Windham, A.E. Yagle, Optimal transformation for correcting partial volume averaging effects in magnetic resonance imaging, *IEEE Trans. Nucl. Sci.* 40 (1993) 1204–1212.
- [9] T.F. Cox, M.A.A. Cox, *Multidimensional Scaling*, second ed., Chapman&Hall, London, 2001.
- [10] J.A. Hartigan, M.A. Wong, A k-means clustering algorithm, *Appl. Stat.* 28 (1979) 100–108.
- [11] K.C. Li, Sliced inverse regression for dimension reduction, *J. Am. Stat. Assoc.* 86 (1991) 316–342.
- [12] C.H. Chen, K.C. Li, Can SIR be as popular as multiple linear regression? *Stat. Sin.* 8 (1998) 289–316.
- [13] K.C. Li, High dimensional data analysis via the SIR/PHD approach, 2000, Lecture notes that are available at (<http://www.stat.ucla.edu/~kcli/>).
- [14] R.D. Cook, X. Yin, Dimension-reduction and visualization in discriminant analysis, *Aust. NZ. J. Stat.* 43 (2001) 147–200.
- [15] C. Cortes, V. Vapnik, Support-vector network, *Mach. Learning* 20 (1995) 273–297.
- [16] C.W. Hsu, C.J. Lin, A comparison of methods for multi-class support vector machines, *IEEE Trans. Neural Networks* 13 (2002) 415–425.
- [17] C.C. Chang, C.J. Lin, LIBSVM: a library for support vector machines (Version 2.33), 2002, Software available at (<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>).
- [18] R. Collobert, S. Bengio, SVMTool: support vector machines for large-scale regression problems, *J. Mach. Learning Res.* 1 (2001) 143–160.
- [19] H.J. Weaver, *Applications of Discrete Continuous Fourier Analysis*, A Wiley-Interscience Publication, New York, 1983.
- [20] W.S. Torgerson, Multidimensional scaling: I. Theory method, *Psychometrika* 17 (1952) 401–419.
- [21] L. Ferre, Determining dimension in sliced inverse regression and related methods, *J. Am. Stat. Assoc.* 93 (1998) 132–140.
- [22] K.C. Li, On principal Hessian directions for data visualization dimension reduction: another application of Stein's lemma, *J. Am. Stat. Assoc.* 87 (1992) 1025–1039.
- [23] C.H. Chen, K.C. Li, Generalization of Fisher's linear discriminant analysis via the approach of sliced inverse regression, *J. Korean Stat. Soc.* 30 (2001) 193–217.
- [24] J.T. Kent, Discussion of Li (1991), *J. Am. Stat. Assoc.* 86 (1991) 336–337.
- [25] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, Berlin, 2001.
- [26] V.N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, 1995.
- [27] J.C. Platt, N. Cristianini, J. Shawe-Taylor, Large margin DAGs for multiclass classification, *Adv. Neural Inf. Process. Syst.* 12 (2000) 547–553.
- [28] H.M. Wu, Kernel sliced inverse regression with applications on classification, Technical Report 2006, Institute of Statistical Science, Academia, Taiwan.
- [29] H.M. Wu, H.H.-S. Lu, Supervised motion segmentation by spatial-frequency analysis and dynamic sliced inverse regression, *Stat. Sin.* 14 (2004) 413–430.

**About the Author**—HAN-MING WU received his B.S. degree in Mathematics from the Tamkang University, Taiwan, in 1995, the M.S. degree in Mathematical Statistics from the National Chung Cheng University, Taiwan, in 1997, and the Ph.D. degree in Statistics from the National Chiao Tung University, Taiwan, in 2003. He is working as a Postdoctoral Fellow in the Institute of Statistical Science, Academia Sinica, Taipei, Taiwan. His research interests include information visualization, bioinformatics, medical image processing, and statistical computing.

**About the Author**—HENRY HORNG-SHING LU received his Ph.D. and M.S. degrees in Statistics from the Cornell University, NY, USA, in 1994 and 1990, respectively, and his B.S. degree in electric engineering from the National Taiwan University, Taiwan, in 1986. He is a Professor in the Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan. His areas of research include statistics, medical images, and bioinformatics.