www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

COMPUTER VISION-BASED COLOR IMAGE SEGMENTATION WITH IMPROVED KERNEL CLUSTERING

Yongqing Wang[1, 2*] and Chunxiang Wang[3]

[1] Department of Computer Science and Applications, Zhengzhou Institute of Aeronautical Industry Management, Zhengzhou 450015, China

[2] Henan aviation economics research center, aviation economic development & aeronautic materials technology collaborative innovation center

[3] Basic Course Department, Henan Polytechnic, Zhengzhou 450046, China

*Email: wyq-yongqing@163.com

Submitted: May 6, 2015 ***Accepted: Jul. 31, 2015*** ***Published: Sep. 1, 2015***

Abstract- Color image segmentation has been widely applied to diverse fields in the past decades for containing more information than gray ones, whose essence is a process of clustering according to the color of pixels. However, traditional clustering methods do not scale well with the number of data, which limits the ability of handling massive data effectively. We developed an improved kernel clustering algorithm for computing the different clusters of given color images in kernel-induced space for image segmentation. Compared to other popular algorithms, it has the competitive performances both on training time and accuracy. The experiments performed on both synthetic and real-world data sets demonstrate the validity of the proposed algorithm.

Index terms*:* Computer vision, Color image segmentation, Kernel clustering, MEB algorithm, Support vector data description.

I. INTRODUCTION

If the eyes are the window to the soul, then the visual ability decides whether or not the window is clear. What we human being can not denying is that, our visual ability is tend to be fragile, fatigue, aging and damageable, and what listed before, to some extent, are on the contrary for computer vision.

Due to the merits of high efficiency, nondestruction, objective and accurate result, and fatigueless working, computer vision has achieved lots of successful applications in more and more fields, such as biological feature recognition, information retrieval, and so on. By utilizing of infrared-ray, ultraviolet-ray, X-ray, ultrasound and other advanced detection techniques, computer vision has the significant advantages in detecting invisible objects and under high dangerous scenes. The various application areas can be summarized but not limited to as follows.

a. Application in industrial detection

Nowadays, computer vision has been successfully applied in the field of industrial detection (seen in Fig. 1), and greatly improves the quality and reliability of the product, which guarantees the speed of production.

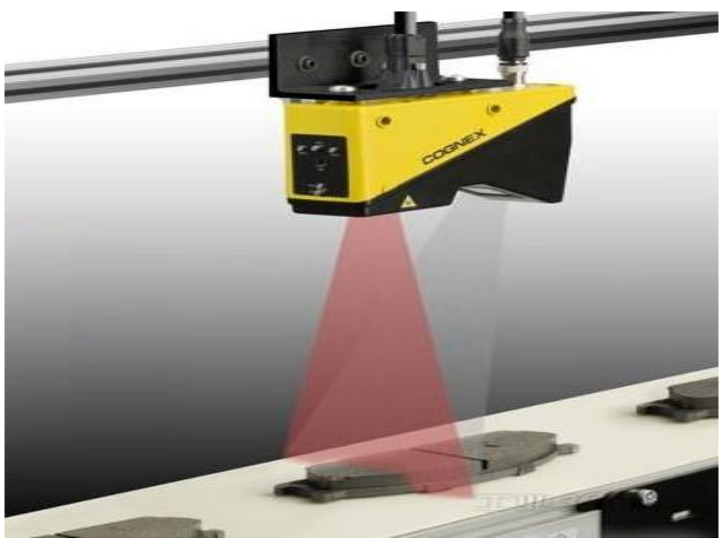

Figure 1. Application of computer vision in industry detection

For example, the quality detection in packaging and printing for products, quality detection for containers, drink filling, and bottle cap sealing in beverage industry, timber wood detection, semiconductor integrated packaging quality detection, coil quality detection, the industrial computed tomography of key mechanical parts, etc. At the customs, the application of X-ray and the computer vision technology can inspect the cargo without opening the package, which can greatly improve the speed of customs clearance, and saves a large amount of manpower and material resources. In the pharmaceutical production line, computer vision technology can be used to test the drug packaging, which can guarantee the package quality of drags.

b. Application in aviation and remote sensing

Computer vision can be used in diverse scenes of aviation and remote sensing (seen in Fig. 2), which can be listed but not limited to the following applications.

Figure 2. Application of computer vision in remote sensing: Surreal Bridges of Google Earth

The reconnaissance, positioning and navigation in military scenes. Automatic cartography, satellite images and topographic map alignment, automatic surveying and mapping. Management of land and resources, such as the management of the forest, water, soil, etc. Synthetic analysis and prediction to weather forecast, automatic environment and fire alarm monitoring. Detection and analysis of astronomy and space objects, transportation and air lines management, etc. Collecting satellite remote sensing images, automatic identifying and classifying the ground targets according to the characteristics of image and graphics topography, etc.

c. Application in biomedicine

In the field of biomedicine, computer vision is used to assist doctors in medical images analysis (seen in Fig. 3), where digital image processing and information fusion technologies can be used for medical imaging data statistics and analysis in the scenes of X-ray perspective, nuclear magnetic resonance and CT images. For example, X-ray images reflect the bone tissue, nuclear magnetic resonance images reflect the organic organization, and doctors often need to consider the relationship between skeleton and organic organization, therefore the digital image processing technique is required to suitable superpose two kinds of images together to facilitate the medical analysis.

Figure 3. Application of computer vision in CT detection

d. Application in military and public security

The application of computer vision in security includes two scenarios, military security and public security.

d.i Military security

The military security scenario includes the cruise missile terrain recognition, the object recognition and tracking (seen in Fig. 4), the radar terrain reconnaissance, the remote control aircraft guidance, the target identifying and guidance, the military alert system and automatic control of artillery, etc.

Figure 4. Application of computer vision in military object tracking

d.ii Public security

The public security scenario includes the fingerprint automatic recognition, the iris feature automatic recognition, the synthesis of criminals' face, the automatic identification of handwriting, portrait, and seal, enhancing image quality to capture emergency in the monitoring system for closed circuit television (seen in Fig. 5), intelligent scheduling in traffic management system, etc.

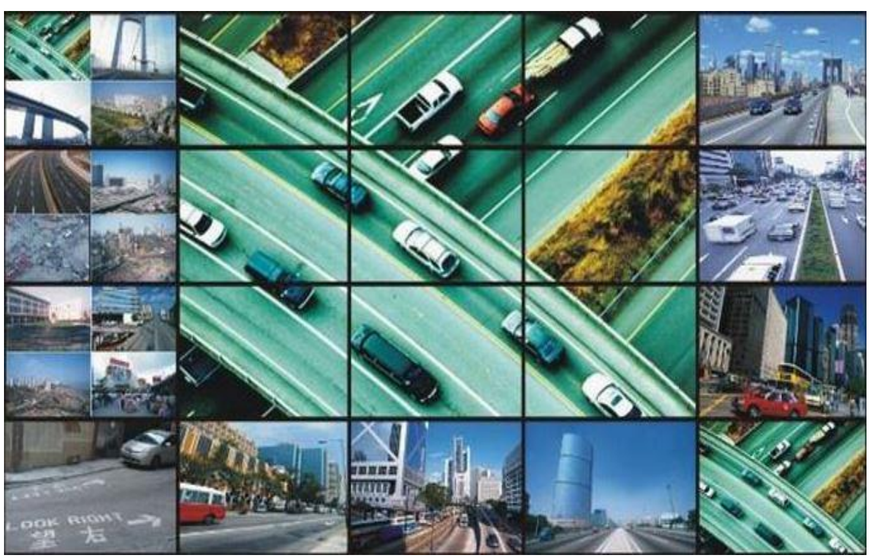

Figure 5. Application of computer vision in traffic monitoring

The applications of computer vision have many other topics, such as image segmentation. Being one of the primal problems in image processing, image segmentation can be divided into two main types according to different strategies, gray image segmentation and color image segmentation. Due to the fact that color images contain more information to describe the real-world more vividly than gray images, along with the rapid increasing of computer processing capacity, color image segmentation has attracted more attention than gray image segmentation during the past few years.

Traditional color image segmentation methods can be implemented with different ideas, such as edge segmentation, region segmentation, segmentation based on threshold, segmentation based on artificial neural network, segmentation based on wavelet, segmentation based on active contour model, segmentation based on genetic algorithm, clustering segmentation, and so on. As the essence of color image segmentation is a kind of clustering according to the color of pixels, large number of clustering segmentation methods appeared in recent years.

But, traditional clustering methods do not scale well with the training sample size, which limits the ability of color image segmentation to handle massive data effectively. Inspired by the idea proposed by Support Vector Data Description, we developed an improved kernel clustering algorithm for computing the different clusters of given color images in kernel-induced space to segment the color images. Compared to other popular algorithms, it has the competitive performances both on training time and accuracy. The experiments performed on both synthetic and real-world data sets demonstrate the validity of the proposed algorithm.

The rest of this paper is organized as follows. Section II provides a review on Color Image Segmentation. Popular clustering methods are introduced in Section III. Section IV proposes the improved Kernel Clustering algorithm in detail. Section V conducts the experimental results on both synthetic and real-world data set, and Section VI ends this paper with a conclusion.

II. COLOR IMAGE SEGMENTATION

Comparing by the principle of segmentation, color image segmentation and gray image segmentation are the same techniques since they are both based on pixel numerical similarity and proximity of input space, except for the transition of the investigation on the pixel attribute and feature extraction technology from one dimension space to high dimension space. With the rapid increasing of computer processing capacity in recent years, color image segmentation has attracted increasingly more attention. According to different strategies adopted in handling methods, color image segmentation can be divided into categories as follows.

a. The method based on neighborhood

According to the same or similar characters of pixels, the method based on neighborhood can connect the neighboring pixels to achieve image segmentation. It includes all sorts of region growing methods, watershed segmentation method, and the Markov Random Field (MRF) method. This method can make a full use of space information and the correlation between pixels. However, some priori information is needed at the same time, such as the seed pixels and a variety of criteria to define the color target boundaries, which are difficult to get in some cases.

a.i Region growing method

Region growing method (including region splitting-and-merging technology), initializes from several seed points or seed regions, according to certain growth standards, discriminates and connects the points of neighborhood pixels, until all pixel points have been connected [1-3]. This method defines three principles in Euclid distance of RGB color space, color identity principle, principle based on the color similarity and space neighborhood, global identity principle based on the color similarity, which leads to the subjectivity in the selection of the corresponding thresholds, and not suitable for image segmentation with shadow area.

a.ii Watershed segmentation method

Watershed segmentation method [4-7] is based on gradient, whose advantage is that it can get one pixel wide, which can result in a closed connected precise outline. However, it is difficult to select a proper label in watershed segmentation method, and improper labeling always leads to an over-segmentation result.

a.iii Markov Random Field method

Markov random field method [8-10] is one of the most commonly used statistical methods in image segmentation, especially in the widespread application of the texture image. Its essence is regarding the color value of each point in image as a random variable with certain probability distribution, where the probability of a pixel point taking some color is determined by its neighborhood other than the global information of the image. Technologies based on random field model can provide more accurate image area domain feature. When confronted with complex image area, or difficulty to divide the image through simple technology, the methods based on random field model will always achieve very good segmentation results. However, it needs a lot of calculation, and the related algorithms are very complicated, so the balance between computing complexity and making good segmentation results is a challenging topic to be further studied.

b. The method based on histogram threshold

The method based on histogram threshold [11-13] uses the valley value between the two adjacent peaks in color histogram as a threshold to segment images. Unlike gray images, color images have three color components, and the histogram is a 3-d array, by which it is more difficult to determine the thresholds. The most used solving method includes adopting three two-dimensional spectrum subsets for RGB space, and choosing three main spectrum subsets for images with more spectrum subsets by the use of principal component transform. Histogram threshold-based method does not need any priori information, and the amount of calculation is small. However, the segmented areas based on color segmentation may be incomplete, and the results for images without obvious peaks are poor.

c. The method based on clustering

The method of color image segmentation based on clustering [14-18], represents the pixels in the image space with the corresponding feature space points, segments the feature space points according to their clustering in feature space, and then maps them back to the original image space to get the segmentation result.

The clustering-based color image segmentation method can be classified into two categories, multi-dimensional color clustering and multi-dimensional extension of histogram threshold, both of which have its own merits and shortcomings. There exists some dependent losses and correlation among multi-dimensional data in multi-dimensional color clustering, but it cannot correctly represent data clustering. However, the multi-dimensional extension of histogram threshold method has high efficiency in calculation; but on the other hand, it cannot represent the color feature space very well. The former has higher computation costs than the latter, but it can represent the color space better.

At present, there are many strategies combining multidimensional threshold segmentation with other methods, such as 3-d histogram growth method, scale spatial clustering method, the dynamic clustering algorithm, etc., which can overcome the phenomenon of excessive segmentation with simple process of classification, and it is easy to implement. However, it has also some disadvantages. Firstly, there is the problem of how to determine the number of colors in color image segmentation. Secondly, the characteristics are usually independent of the images, and there is the question of how to select features in order to get satisfactory separation results. Thirdly, there is no good use of spatial information, which is very useful for image segmentation.

d. The method combined with other theories

Except for the methods mentioned above, there are some methods combined with other theories, which can be listed as follows [19-24]. The color image segmentation technology based on fuzzy set theory, the color image segmentation technology based on wavelet analysis, the color image segmentation technology based on neural network, and the method based on physical model, all of which are suitable for some certain kind of application, and a unified framework is to be further studied.

III. CLUSTERING METHODS

Being one kind of common tool of data analysis and unsupervised machine learning methods, clustering aims at dividing the data set into several classes (or clusters), keeping the maximum similarity between the data of each same class, and the maximum difference between the data of each pair of different class [25] (seen in Fig. 6). According to the basic ideas the clustering algorithms adopt, the popular methods can be roughly divided into the types listed as follows [26].

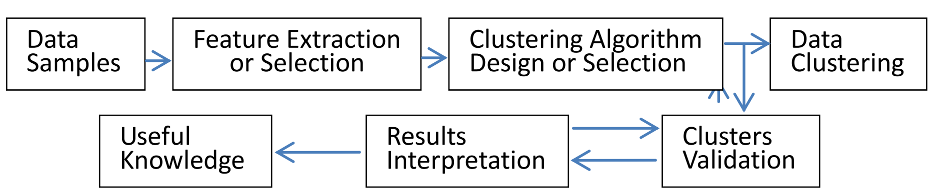

Figure 6. The flow diagram of clustering

a. Hierarchical clustering

Hierarchical clustering (HC) algorithms organize data into a hierarchical structure according to the proximity matrix. The results of HC are usually depicted by a binary tree or dendrogram. The root node of the dendrogram represents the whole data set and each leaf node is regarded as a data object. The intermediate nodes, thus, describe the extent that the objects are proximal to each other; and the height of the dendrogram usually expresses the distance between each pair of objects or clusters, or an object and a cluster. The ultimate clustering results can be obtained by cutting the dendrogram at different levels. This representation provides very informative descriptions and visualization for the potential data clustering structures, especially when real hierarchical relations exist in the data, like the data from evolutionary research on different species of organisms. HC algorithms are mainly classified as agglomerative methods and divisive methods.

Agglomerative clustering starts with *N* clusters and each of them includes exactly one object. A series of merge operations are then followed out that finally lead all objects to the same group. Divisive clustering proceeds in an opposite way. In the beginning, the entire data set belongs to a cluster and a procedure successively divides it until all clusters are singleton clusters.

b. Partitional clustering

In contrast to hierarchical clustering, which yields a successive level of clusters by iterative fusions or divisions, partitional clustering assigns a set of objects into *K* clusters with no hierarchical structure. In principle, the optimal partition, based on some specific criterion, can be found by enumerating all possibilities. But this brute force method is infeasible in practice, due to the expensive computation.

b.i Squared error-based clustering

One of the important factors in partitional clustering is the criterion function. The sum of squared error function is one of the most widely used criteria. The *K*-means algorithm is the best-known squared error-based clustering algorithm, whose algorithm is very simple and can be easily implemented in solving many practical problems. It can work very well for compact and hyper spherical clusters. Parallel techniques for *K*-means are developed that can largely accelerate the algorithm. The drawbacks of *K*-means are also well studied, and as a result, many variants of *K*-

means have appeared in order to overcome these obstacles. We summarize some of the major disadvantages with the proposed improvement in the following.

1) There is no efficient and universal method for identifying the initial partitions and the number of clusters. The convergence centroids vary with different initial points. A general strategy for the problem is to run the algorithm many times with random initial partitions.

2) The iteratively optimal procedure of *K*-means method cannot guarantee the convergence to a global optimum.

3) *K*-means is sensitive to outliers and noise. Even if an object is quite far away from the cluster centroid, it is still forced into a cluster and, thus, distorts the cluster shapes.

4) The definition of "means" limits the application only to numerical variables. The *K*-medoids algorithm mentioned previously is a natural choice, when the computation of means is unavailable, since the medoids do not need any computation and always exist.

b.ii Mixture densities-based clustering

In the probabilistic view, data objects are assumed to be generated according to several probability distributions. Data points in different clusters were generated by different probability distributions. They can be derived from different types of density functions (e.g., multivariate Gaussian or *t*-distribution), or the same families, but with different parameters. If the distributions are known, finding the clusters of a given data set is equivalent to estimating the parameters of several underlying models.

As long as the parameter vector is decided, the posterior probability for assigning a data point to a cluster can be easily calculated with Bayes's theorem. Here, the mixtures can be constructed with any types of components, but more commonly, multivariate Gaussian densities are used due to its complete theory and analytical tractability. Maximum likelihood (ML) estimation is an important statistical approach for parameter estimation and it considers the best estimate as the one that maximizes the probability of generating all the observations.

Unfortunately, since the solutions of the likelihood equations cannot be obtained analytically in most circumstances, iteratively suboptimal approaches are required to approximate the ML estimates. Among these methods, the expectation-maximization (EM) algorithm is the most popular.

The major disadvantages for EM algorithm are the sensitivity to the selection of initial parameters, the effect of a singular covariance matrix, the possibility of convergence to a local optimum, and the slow convergence rate.

c. Graph theory-based clustering

The concepts and properties of graph theory make it very convenient to describe clustering problems by means of graphs. Nodes *V* of a weighted graph *G* correspond to data points in the pattern space and edges E reflect the proximities between each pair of data points. Both the single linkage HC and the complete linkage HC can be described on the basis of the threshold graph. Single linkage clustering is equivalent to seeking maximally connected subgraphs (components) while complete linkage clustering corresponds to finding maximally complete subgraphs (cliques). Graph theory can be used for hierarchical or nonhierarchical clusters.

d. Combinatorial search techniques-based clustering

The basic object of search techniques is to find the global or approximate global optimum for combinatorial optimization problems, which usually have NP-hard complexity and need to search an exponentially large solution space. Clustering can be regarded as a category of optimization problems. Given a set of *N* data points, clustering algorithms aim to organize them into *K* subsets. Even for small *N* and *K*, the computational complexity is extremely expensive, not to mention the large-scale clustering problems frequently encountered in recent decades. Simple local search techniques, like hill-climbing algorithms, are utilized to find the partitions, but they are easily stuck in local minima and therefore cannot guarantee optimality. More complex search methods, evolutionary algorithms (EAs) and Tabu search (TS) are known as stochastic optimization methods, while deterministic annealing (DA) is the most typical deterministic search technique, can explore the solution space more flexibly and efficiently.

e. Fuzzy Clustering

So far, the clustering techniques discussed before are referred to hard or crisp clustering, which means that each object is assigned to only one cluster. For fuzzy clustering, this restriction is relaxed, and the object can belong to all of the clusters with a certain degree of membership. This is particularly useful when the boundaries among the clusters are not well separated and ambiguous. Moreover, the memberships may help us discover more sophisticated relations between a given object and the disclosed clusters. The standard Fuzzy Clustering Method (FCM) alternates the calculation of the membership and prototype matrix, which causes a computational burden for large-scale data sets. Kolen and Hutcheson accelerated the computation by combining updates of the two matrices. Hung and Yang proposed a method to reduce computational time by identifying more accurate cluster centers. FCM variants were also developed to deal with other data types, such as symbolic data and data with missing values.

f. Neural networks-based clustering

In neural networks-based clustering, input patterns are fully connected to all neurons via adaptable weights, and during the training process, neighboring input patterns are projected into the lattice, corresponding to adjacent neurons. Neural networks-based clustering has been dominated by SOFMs and adaptive resonance theory (ART). In competitive neural networks, active neurons reinforce their neighborhood within certain regions, while suppressing the activities of other neurons. In addition to these, many other neural network architectures are developed for clustering. Most of these architectures utilize prototype vectors to represent clusters, e.g., cluster detection and labeling network.

Except for the clustering methods mentioned above, there are many other methods appear recently. The research on the clustering algorithm is deepening presently, and the kernel clustering and spectral clustering are two methods that have attracted much attention in recent years [27].

IV. IMPROVED KERNEL CLUSERING METHODS

The main idea of kernel clustering method is adopting a nonlinear mapping φ, such that the data points in input space can be mapped into a high-dimensional feature space, selecting appropriate Mercer kernel function instead of nonlinear mapping of the inner product to cluster in the feature space. The kernel clustering method is universal, and has great improvement to the classical clustering methods. The adopted nonlinear mapping can increase the linear separable probability on input data points, which can achieve more accurate clustering, and faster convergence speed, as well. Under the condition the classical clustering algorithms fail, the kernel clustering algorithms can always work. The kernel trick idea in kernel clustering method can be illustrated in Fig. 7 below, where the left is the original input space, and the right is the kernel-induced space.

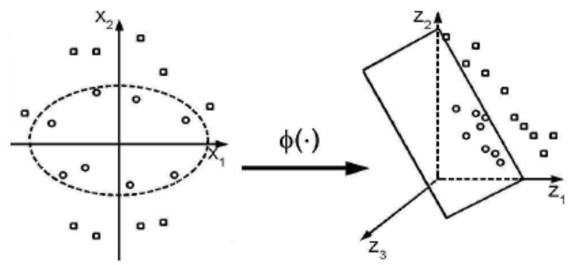

Figure 7. The kernel trick: Nonlinear problems are linear in kernel-induced space

Kernel trick can link many different kernel methods together, one of the most well-known is Support Vector Data Description (SVDD), where SVM and MEB were firstly connected.

a. Support vector data description

The idea of SVDD [29] can be formulated as follows. Formulating Binary SVM as a QP to maximize the margin between two classes, and the consequent generalization ability is always better than the other machine learning methods.

Given a training data sets $S=\{(x_i, y_i) \mid i=1,\ldots,m\}$, where $x_i \in R^d$ and $y_i \in \{+1,-1\}$, the primal for the Binary SVM problem can be formulated as

$$\begin{aligned} \min_{w,\rho,b,\xi_i} \quad & ||w||^2 + b^2 - 2\rho + C\sum_{i=1}^{m}\xi_i^2 \\ s.t. \quad & y_i(w'\phi(x_i)+b) \geq \rho - \xi_i,\ i=1,\ldots,m. \end{aligned} \tag{1}$$

The corresponding dual is

$$\begin{aligned} \min_{\alpha_i} \quad & \sum_{i,j=1}^{m}\alpha_i\alpha_j\left(y_i y_j k(x_i, x_j) + y_i y_j + \frac{\delta_{ij}}{C}\right) \\ s.t. \quad & \sum_{i=1}^{m}\alpha_i = 1,\ \ \alpha_i \geq 0,\ \ i=1,\ldots,m, \end{aligned} \tag{2}$$

Where δ_{ij} is the Kronecker delta function, defined as following

$$\delta_{ij} = \begin{cases} 1, & if\ i=j, \\ 0, & if\ i \neq j. \end{cases} \tag{3}$$

We denote the pair (x_i, y_i) as z_i to simplify the notation. Introducing a modified feature map $\tilde{\phi}(z_i) = [y_i\phi'(x_i)\ \ y_i\ \ \frac{e'_i}{\sqrt{C}}]'$ and the associated kernel function $\tilde{k}(z_i, z_j) = y_i y_j k(x_i, x_j) + y_i y_j + \frac{\delta_{ij}}{C}$, then the dual of Binary SVM with form (2) can be rewritten as

$$\begin{aligned} \min_{\alpha_i} \quad & \sum_{i,j=1}^{m} \alpha_i \alpha_j \tilde{k}(z_i, z_j) \\ s.t. \quad & \sum_{i=1}^{m} \alpha_i = 1,\ \alpha_i \geq 0,\ \ i = 1,\ldots,m. \end{aligned} \tag{4}$$

b. Support vector clustering

Support Vector Clustering (SVC) belongs to the method of kernel clustering, whose foundational tool for clustering is Support Vector Machine (SVM) [28]. Based on the Support Vector Domain Description (SVDD) algorithm [29], Ben-Hur (2001) proposed an unsupervised nonparametric clustering algorithm, called SVC [30], whose aim is to find a set of contours used as the cluster boundaries in the original data space. These contours can be formed by mapping back the smallest enclosing sphere in the transformed feature space. RBF is chosen in this algorithm, and, by adjusting the width parameter of RBF, SVC can form either agglomerative or divisive hierarchical clusters. When some points are allowed to lie outside the hypersphere, SVC can deal with outliers effectively.

Kernel-based clustering algorithms have many advantages.

1) It is more possible to obtain a linearly separable hyperplane in the high-dimensional, or even infinite feature space.

2) They can form arbitrary clustering shapes other than hyperellipsoid and hypersphere.

3) Kernel-based clustering algorithms, like SVC, have the capability of dealing with noise and outliers.

4) For SVC, there is no requirement for prior knowledge to determine the system topological structure.

But there is still a bottleneck of weak scalability with the number of training sample data size in SVC algorithm, so many new SVC algorithms are designed to improve the computing efficiency [30-34].

c. Improved kernel clustering

In this subsection, we proposed an improved kernel clustering (IKC) algorithm in detail, the concise conclusions about the time and space complexities of the proposed algorithm were presented, as well.

The detailed procedure of IKC was given in Table 1 below.

c.i IKC algorithm in detail

Table 1: The procedure of IKC Algorithm

	IKC Algorithm
1:	Given $\epsilon > 0$, pick any $\varphi(\mathbf{p}) \in S$, find $\varphi(\mathbf{q}) \in S$ that is furthest away from $\varphi(\mathbf{p})$, $\mathbf{c}_0 \leftarrow \frac{1}{2}(\varphi(\mathbf{p}) + \varphi(\mathbf{q}))$, $r_0 \leftarrow \frac{1}{2}\|\varphi(\mathbf{p}) - \varphi(\mathbf{q})\|$.
2:	Terminate if no $\varphi(\mathbf{z})$ falls outside $B(\mathbf{c}_t, (1+\epsilon)r_t)$. Otherwise, find $\varphi(\mathbf{z}_t)$ that is furthest away from $\mathbf{c}_t$.
3:	Find the smallest update to the center such that $B(\mathbf{c}_{t+1}, (1+\epsilon)r_t)$ touches $\varphi(\mathbf{z}_t)$.
4:	Increment t by 1 and go back to Step 2.

The proposed iterative IKC algorithm is shown in Table 1 above. There are two strategies to ensure the high efficiency. Firstly, the update in Step 3 can be performed efficiently without the use of any numerical optimization solver. Secondly, the radius in IKC is asymptotically expansive and converges to the optimum in $O(\frac{1}{\varepsilon})$ iterations within any precision. The efficient update of the t[th] iteration is shown in Fig. 8.

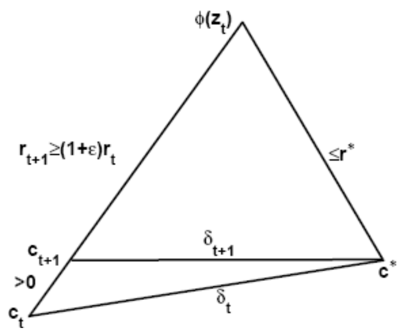

Figure 8. Update of c_t at the t^{th} iteration

Mathematically analyzing, IKC seeks for the c_{t+1}, r_{t+1} such that

$$\begin{aligned} &\min_{c_{t+1}, r_{t+1}} \ ||c_{t+1} - c_t||^2 \\ &s.t. \ ||c_{t+1} - \varphi(z_t)||^2 \le r_{t+1}{}^2, \\ &\quad (1+\varepsilon) r_t \le r_{t+1}. \end{aligned} \quad (5)$$

From the KKT conditions on form (5) we can get the new center is $c_{t+1} = \beta_t c_t + (1-\beta_t)\varphi(z_t)$, where $\beta_t = \frac{(1+\varepsilon)r_t}{\|c_t - \varphi(z_t)\|} (\ge 0)$, which is a convex combination of c_t and $\varphi(z_t)$. Consequently, for any $t > 0$, c_t is always a convex combination of c_0 and $S_t = \{\varphi(z_i)\}_{i=1}^{t}$, i.e., $c_t = \sum_{i=0}^{t} \alpha_i \varphi(z_i)$, where $\sum_{i=0}^{t} \alpha_i = 1,\ \alpha_i \ge 0,\ \varphi(z_0) = c_0$.

c.ii Performance analysis of IKC algorithm

In this section we conclude that the iterative IKC algorithm converges to the optimum within any precision, and the time and space complexities are superior to algorithms of CVM and SVC (seen in Fig. 9).

Proposition 1. IKC algorithm obtains an $(1+\varepsilon)$-approximation of MEB(S) in $O(\frac{1}{\varepsilon})$ iterations.

Proposition 2. Assume that the IKC algorithm terminates at the τ^{th} iteration with the solution (r_τ, c_τ), then $\|c_\tau - c^*\| \le \sqrt{\varepsilon(2+\varepsilon)}r^*$, where (r^*, c^*) denotes the optimum of MEB(S).

Proposition 3. The time complexity of IKC algorithm is $O(\frac{m}{\varepsilon^2}+\frac{1}{\varepsilon^3})$, which is linear in m for a fixed ε.

Proposition 4. The space complexity of IKC algorithm is $O(\frac{1}{\varepsilon^2})$, which is independent of m for a fixed ε.

The detailed proofs of these theorems are omitted here for conciseness, interested readers can refer to Wang [35].

Algorithm	Iterations	Space complexity	Time complexity
CVM	$O(\frac{1}{\epsilon})$	$O(\frac{1}{\epsilon^2})$	$O(\frac{m}{\epsilon^2}+\frac{1}{\epsilon^4})$
SVC	$O(\frac{1}{\epsilon^2})$	$O(\frac{1}{\epsilon^2})$	$O(\frac{1}{\epsilon^4})$
IKC	$O(\frac{1}{\epsilon})$	$O(\frac{1}{\epsilon^2})$	$O(\frac{m}{\epsilon^2}+\frac{1}{\epsilon^3})$

Figure 9. Mathematical comparison on complexities of different algorithms

From the mathematical comparison in Fig. 9 above on the indexes of iterations, space and time complexities, we can conclude that algorithm IKC is of the better performance than the others.

V. EXPERIMENTAL RESULTS

We conduct the experiments on both synthetic and real-world data set to prove the validity and efficiency of the proposed IKC algorithm. All the experiments were done on an AMD Sempron (tm) 2500+ 1.41GHz PC with 1GB RAM, and the software package we utilized is Matlab 7.0 toolbox.

Firstly, we compare the kernel methods of Core Vector Machine (CVM), Support Vector Clustering (SVC), and the proposed Improved Kernel Clustering (IKC) algorithm on six synthetic data sets, which are generated randomly and follow the uniform distribution.

a. Experiment on synthetic data set

The details of data sets used in this section are listed in Table 2 below.

Table 2: Details of synthetic data sets

Data set	**Sy. 1**	**Sy. 2**	**Sy. 3**	**Sy. 4**	**Sy. 5**	**Sy. 6**
# Class	4	4	4	4	4	4
# Dim.	2	2	2	2	2	2
# Point	20	100	200	1000	2000	10000

We conduct the experiments with different ε on these data sets for all the algorithms to compare the performances. Core vectors' number, support vectors' number and raining time for all the algorithms, vary with data size on the synthetic data under the best choice of ε are given in Fig. 10-Fig. 12. We can see that the proposed IKC algorithm is of the smallest core vectors' number, support vectors' number, and the shortest training time.

Generally speaking, we can conclude that, compared to methods of CVM and SVC, the proposed IKC is of the shortest training time, the highest accuracy and the smallest core vectors' number.

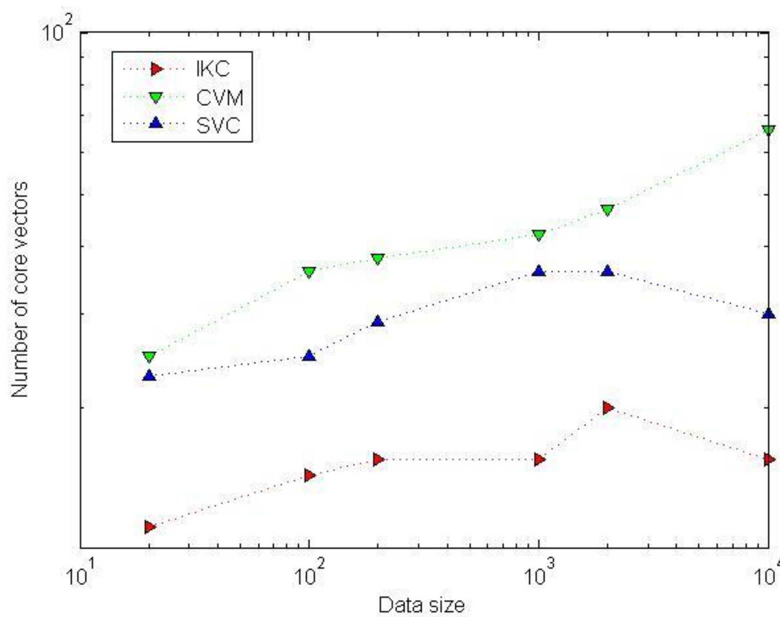

Figure 10. Core vector's number vary with data size for different algorithms under the best ε

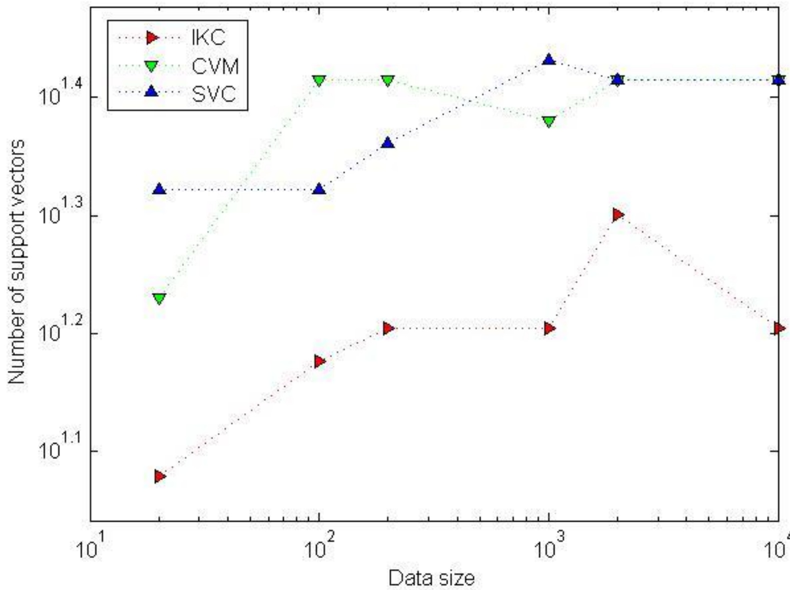

Figure 11. Support vector's number vary with data size for different algorithms under the best ε

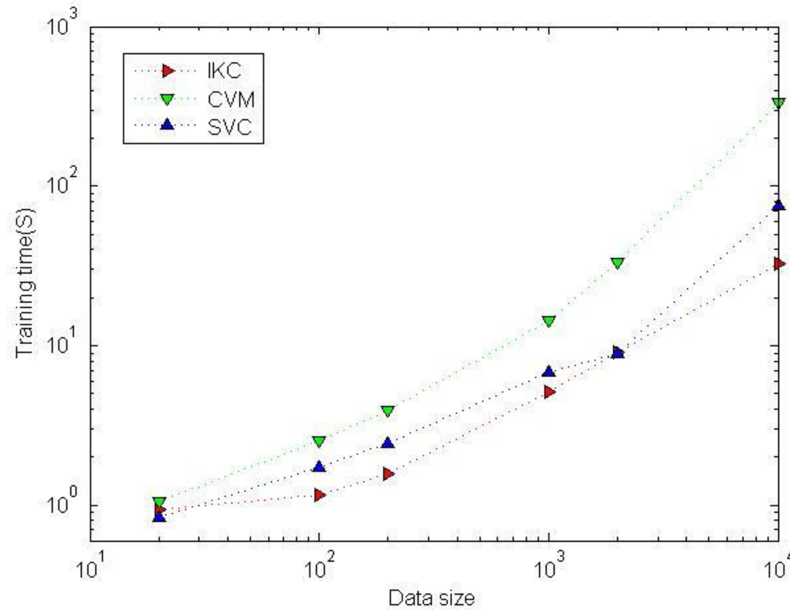

Figure 12. Training time vary with data size for different algorithms under the best ε

b. Experiment on real-world color images

Utilizing the proposed IKC algorithm, we can handle the color image segmentation problem. All of the original images to be segmented are chosen from internet by personal interest, including three helicopter images, each of which has different background with different noisy level. The segmentation results are demonstrated in Fig. 13-Fig. 15, where the left images stand for the original color images, the middle ones stand for the segmented color images under the mean-shift algorithm, and the right images stand for the segmented color images under the IKC algorithm we proposed.

From the comparison results we can see that the proposed IKC algorithm can achieve competitive performances on color image segmentation.

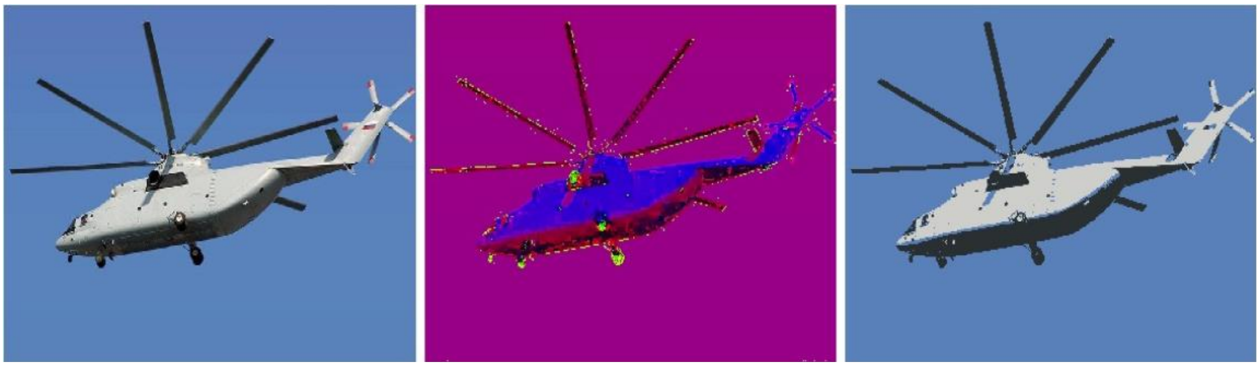

Figure 13. Original color image (left) vs. segmented image under mean-shift (middle) vs. segmented image under IKC (right) for helicopter with low noisy background

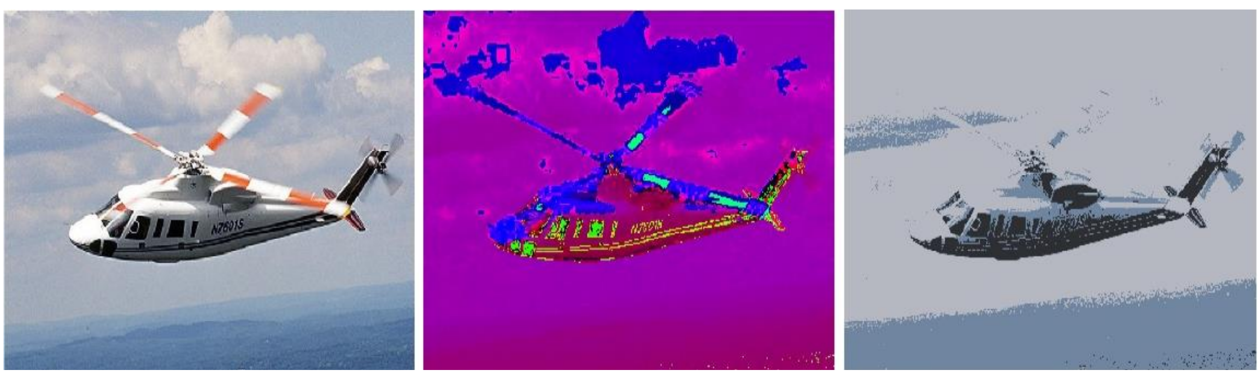

Figure 14. Original color image (left) vs. segmented image under mean-shift (middle) vs. segmented image under IKC (right) for helicopter with medium noisy background

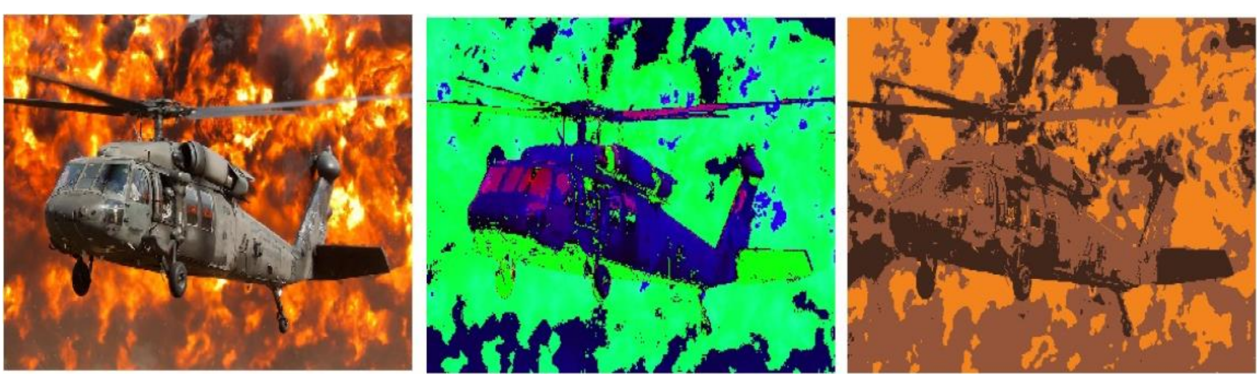

Figure 15. Original color image (left) vs. segmented image under mean-shift (middle) vs. segmented image under IKC (right) for helicopter with high noisy background

VI. CONCLUSIONS

We develop an improved kernel clustering algorithm to handle color image segmentation problem in this article. We conclude with some proved propositions that the proposed IKC algorithm has time complexity of $O(\frac{m}{\varepsilon^2}+\frac{1}{\varepsilon^3})$, which is linear in the number of training samples m for a fixed ε, and space complexity of $O(\frac{1}{\varepsilon^2})$, which is independent of m for a fixed ε. Compared to CVM and SVC algorithms, it has the competitive performances in both training time and accuracy. Besides, by use of the proposed IKC algorithm, we can achieve a fast process to effective handle color image segmentation problem. Experiments on both synthetic and real-world data sets demonstrate the validity of the proposed algorithm.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (Grant No. 41001235), the Project of Henan Provincial Audit Office (Grant No. 20120922), the Scientific & Technological Research Key Project of Henan Provincial Education Office (Grant No. 13A520404), the Scientific & Technological Project of Henan Provincial Science & Technology Office (Grant No. 132102210468), the Project of Science and Technology Bureau of Zhengzhou (Grant No. 20120435 and 20130713).

The authors are grateful to the anonymous referees for their valuable comments and suggestions to improve the presentation of this paper.

REFERENCES

[1] M. Celenk, “A color clustering technique for image segmentation”, Computer Vision, Graphics, and Image Processing, Vol. 52, No. 2, 1990, pp. 145-170.

[2] A. Tremeau, N. Borel, “A region growing and merging algorithm to color segmentation”, Pattern Recognition, Vol. 30, No. 7, 1997, pp. 1191-1203.

[3] H. D. Cheng, Y. Sun, “A hierarchical approach to color image segmentation using homogeneity”, IEEE Trans. on Image Processing, Vol. 9, No. 12, 2000, pp. 2071-2082.

[4] O. Lezoray, H. Cardot, “Cooperation of color pixel classification schemes and color watershed: a study for microscopic images”, IEEE Transactions on Image Processing, Vol. 11, No. 7, 2002, pp. 783-789.

[5] L. Shafarenko, M. Petrou, J. Kittler, “Automatic watershed segmentation of randomly textured color images”, IEEE Trans. on Image Processing, Vol. 6, No. 11, 1997, pp. 1530-1544.

[6] A. Shiji, N. Hamada, “Color image segmentation method using watershed algorithm and contour information”, International Conference on Image Processing, 1999, pp. 305-309.

[7] L. H. Ma, Y. Zhang, J. P. Deng, “A target segmentation algorithm based on opening-closing binary marker on watersheds and texture merging”, Journal of Image and Graphics, Vol. 8, 2003, No 1, pp. 77-83.

[8] G. Celeux, F. Forbes, N. Peyrard, “EM procedures using mean field-like approximations for Markov model-based image segmentation”, Pattern Recognition, Vol. 36, No. 1, 2003, pp.131-144.

[9] Y. Zhang, M. Brady, S. Smith, “Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm”, IEEE Transactions on Medical Imaging, Vol. 20, No. 1, 2001, pp. 45-57.

[10] C. L. Huang, T. Y. Cheng, C. C. Chen, “Color images segmentation using scale space filter and Markov random field”, Pattern Recognition, Vol. 25, No. 10, 1992, pp. 1217-1229.

[11] G. D. Guo, S. Yu, S. D. Ma, “Unsupervised segmentation of color images”, Proc. ICIP, 1998, pp. 299.

[12] F. Kurugollu, B. Sankur, “Color image segmentation using histogram multi-thresholding and fusion”, Image and Vision Computing, Vol. 19, No. 13, 2001, pp. 915-928.

[13] C. Q. Liu, H. Cheng, “An efficient clustering method for color image segmentation”, Pattern Recognition and Artificial Intelligence, Vol. 8, No. A01, 1995, pp. 133-138.

[14] D. Carevic, T. Caelli, “Region-Based Coding of Color Images Using Karhunen-Loeve Transform”, Computer Vision, Graphics, and Image Processing, Vol. 59, No. 1, 1997, pp. 27-38.

[15] X. W. Wang, L. S. Shen, B. G. Wei, “The focus segmentation of color ophthalmologic image based on modified K-means clustering and mathematical morphology”, Chinese Journal of Biomedical Engineering, Vol. 5, 2002.

[16] N. Zahid, M. Limouri, A. Esseaid, “A new cluster-validity for fuzzy clustering”, Pattern Recognition, Vol. 32, No. 7, 1999, pp. 1089-1097.

[17] K. Y. Lin, L. H. Xu, J. H. Wu, "A fast fuzzy C-means clustering for color image segmentation" , Journal of Image and Graphics, Vol. 9, No. 2, 2004.

[18] T. Q. Chen, Y. Lu, "Color image segmentation-An innovative approach", Pattern Recognition, Vol. 35, No. 2, 2002, pp. 395-405.

[19] M. Sammouda, R. Sammouda, N. Niki, "Segmentation and analysis of liver cancer pathological color image based on artificial neural networks", ICIP'99, Vol. 3, 1999, pp. 392-396.

[20] M. Sammouda, R. Sammouda, "Improving the performance of Hopfield neural network to segment pathological liver color images", Proceedings of the 17th International Congress and Exhibition, London, June 25-28, 2003.

[21] S. H. Ong, N. C. Yeo, K. H. Lee, "Segmentation of color images using a two-stage self-organizing network", Image and Vision Computing, Vol. 4, 2002.

[22] N. Papamarkos, C. Strouthopoulous, and L. Andreadis, "Multithresholding of color and gray-level images through a neural network technique", Image and Vision Computing, Vol. 18, No. 3, 2000, pp. 213-222.

[23] N. K. Suryadevara and S. C. Mukhopadhyay, "Wireless Sensor Network Based Home Monitoring System for Wellness Determination of Elderly", IEEE Sensors Journal, Vol. 12, No. 6, June 2012, pp. 1965-1972.

[24] N. K. Suryadevara and S. C. Mukhopadhyay, "Determining Wellness Through An Ambient Assisted Living Environment", IEEE Intelligent Systems, May/June 2014, pp. 30-37.

[25] Rui Xu and D. Wunsch, "Survey of Clustering Algorithms", IEEE Transaction on Neural Networks, Vol. 16, No. 3, 2005, pp. 645-678.

[26] J. Han and M. Kamber, Data Mining: Concepts and Techniques, Second Edition, Morgan Kaufmann, San Francisco, 2006.

[27] M. Filippone, F. Camastra, and F. Masulli, "A Survey of Kernel and Spectral Methods for Clustering", Pattern Recognition, Vol. 41, No. 1, 2008, pp. 176-190.

[28] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, Vol. 2, No. 2, 1998, pp. 121-167.

[29] D. M. J. Tax, R. P. W. Duin, "Support Vector Domain Description", Pattern Recognition Letters, Vol. 20, No. 11-13, 1999, pp. 1191-1199.

[30] A. Ben-Hur, D. Horn, H. T. Siegelmann, V. Vapnik, "Support Vector Clustering", Journal of Machine Learning Research, Vol. 2, No. 12, 2001, pp. 125-137.

[31] B. Scholkopf, R. Williamson, A. Smola, J. Shawe-Taylor, J. Platt, “Support Vector Method for Novelty Detection”, Advances in Neural Information Processing System, Vol. 12, 2000, pp. 582-588.

[32] N. K. Suryadevara, S. C. Mukhopadhyay, R. Wang, R. K. Rayudu, “Forecasting the behavior of an elderly using wireless sensors data in a smart home”, Engineering Applications of Artificial Intelligence, Vol. 26, No. 10, 2013, pp. 2641-2652.

[33] G. Sen Gupta, S. C. Mukhopadhyay and M. Finnie, “Wi-Fi Based Control of a Robotic Arm with Remote Vision”, Proceedings of 2009 IEEE I2MTC Conference, Singapore, May 5-7, 2009, pp. 557-562.

[34] Yongqing Wang, “NEW INTELLIGENT CLASSIFICATION METHOD BASED ON IMPROVED MEB ALGORITHM”, International Journal on Smart Sensing and Intelligent Systems, Vol. 7, No. 1, 2014, pp. 72-95.

[35] Yongqing Wang, “FACE RECOGNITION BASED ON IMPROVED SUPPORT VECTOR CLUSTERING”, International Journal on Smart Sensing and Intelligent Systems, Vol. 7, No. 4, 2014, pp. 1807-1829.

[36] Yongqing Wang, “Simpler Minimum Enclosing Ball: Fast Approximate MEB Algorithm for Extentive Kernel Methods”, CCDC’08, 2008, pp. 3492-3497.