A New Approach for Remote Sensing Image Sample Selection Based on Convex Theory

Xin Pan[*], Hongbin Sun
Changchun Institute of Technology, Changchun ,China

Abstract—**Advancements in remote sensing technology have led to improvements in the acquisition of land cover information. The extraction of accurate and timely knowledge about land cover from remote sensing imagery largely depends on the classification techniques used. Support vector machine has been receiving considerable attention as a promising method for classifying remote sensing imagery. However, the support vector machine learning process typically requires a large memory and significant computation time for treating a large sample set, in which some of the samples might be redundant and useless for the support vector machine model training. Therefore, higher-quality and fewer samples from the sample selection should be utilized for support vector machine-based remote sensing classification. A convex theory-based remote sensing sample selection algorithm for support vector machine classifiers is developed in this work. A Landsat-5 Thematic Mapper imagery acquired on August 31, 2009 (orbit number 113/27) is adopted in our experiments. The study area's land cover/use was divided into five categories. Using the region of interest tool, we select samples from the image of the study area, with each category consisting of 1000 independent pixels. Results show that for most cases, our method can achieve higher classification accuracy than random sample selection method.**

Index Terms—**Remote Sensing, Classification, Sample Selection, SVM**

I. INTRODUCTION

Land cover information has been identified as one of the crucial data components for many aspects of global change studies and environmental applications. Remote sensing technology can help obtain land cover information in an easy and timely manner. The remote sensing image classification process is illustrated in Figure 1:
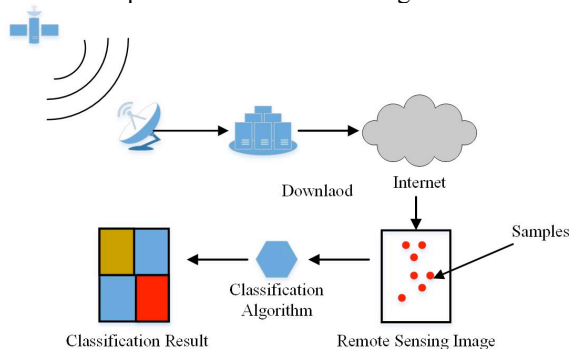

Figure 1.   The process of remote sensing image classification

As shown in Figure 1, the remote sensing satellite collects earth surface images and transmits them to the data center. By specifying the paths and rows, users can download images from the specified location. Classification algorithm analyzes samples with selected pixels and obtains a remote sensing image classification result. Recently, Support Vector Machine (SVM) has been received increasing attention in the study of remote sensing classification [1, 2]. One limitation of SVM is, however, that its training stage takes up large memory and significant computation time, especially in the case of large sample sizes, where some samples may not even be useful for training [3]. Hence, sample selection (i.e. to select the most important samples) plays an important role. Plenty of work for sample selection has been done, including for example based on clustering methods [4,5], Mahalanobis distance [6], β-skeleton and Hausdorff-distance [7,8], and the information theory [9,10]. Although much research progress has been achieved, problems still remain. For a given sample set in a particular application, the majority of existing studies have focused mainly on the acceleration of training speed by minimizing the size of the training sample set; no study has considered the selection of samples with a user specified percentage.

The classification model obtained by SVM is a hyper-plane that maximizes the width of the margin between the classes while minimizing the margin of errors [11,12]. Convex optimization theory was applied in the algorithm to train and find a hyper-plane [13]. This training process, in geometric interpretation, is equivalent to finding the nearest points among convex hulls in Hilbert spaces [14,15].The aforementioned research shows that the position of a sample relative to a convex hull (the geometric interpretation of SVM) can play an important role in classification, specifically for identifying the relationship between training samples and SVM classification results. A convex theory-based remote sensing sample selection algorithm (CTRSSSA) for support vector machine classifiers is developed in this work. A Landsat-5 Thematic Mapper imagery acquired on August 31, 2009 (orbit number 113/27) is adopted in our experiments. The study area's land cover/use was divided into five categories. Using the region of interest tool, we select samples from the image of the study area, with each category consisting of 1000 independent pixels. Results show that for most cases, our method can achieve higher

classification accuracy than random sample selection method.

II. SUPPORT VECTOR MACHINE AND ITS GEOMETRIC INTERPRETATION

Support vector machine (SVM) is a supervised classifier which aims to find hyper-planes that separate the dataset with a maximum margin [12]. Given a set of labeled data (x_1, y_1), ..., (x_n, y_n), where x_i is a multi-dimensional sample vector and $y_i \in \{1, -1\}$ is the class label, the optimization problem associated to the algorithm of SVM can be written as follows [13]:

$$\min W(a) = -\sum_{i=1}^{n} a_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j a_i a_j K(x_i, x_j) \qquad (1)$$

$$s.t. \sum_{i=1}^{n} a_i y_i = 0, \quad 0 \le a_i \le B, \quad i = 1,...,n$$

This is a convex quadratic programming equation $B=\infty$ gives the hard margin case and $B<\infty$ presents the soft margin case. $K(x_i, x_j)$ is the kernel function and the dimension of kernel matrix $K=K(x_i, x_j)$ is equal to the square of the number of training samples, thus more time and computer memory in SVM are needed to train a model when the number of samples increases. The process of finding hyper-plane in SVM training is equivalent to finding the nearest points among convex hulls or reduced convex hulls [14,15].

III. CONVEX THEORY AND SAMPLE SELECTION ALGORITHM

A. Convex theory and distance in Hilbert space

As mentioned above, the position of a sample in a convex can be used in determining the sample's importance for SVM training. Here, convex theory is introduced to find a sample set's convex hull, and based on which, any sample's distance from the convex center in the Hilbert space can be calculated. This distance can be a criterion for sample evaluation and selection.

Definition 1: The set $C \subset R^n$ is convex if [15]

$$ax + (1-a)y \in C, \qquad \forall x, y \in C, a \in [0,1] \qquad (2)$$

Proposition 1: Let C be a nonempty closed convex subset of R^n, and let z be a vector in R^n. There exists a unique vector that minimizes $\|z - x\|$ over $x \subseteq C$, called the projection of z on C. Furthermore, a vector $x*$ is the projection of z on C if and only if

$$(z - x*)'(x - x*) \le 0, \qquad \forall x \in C \qquad (3)$$

Proposition 2: The distance of a point $z \in R^n$ to a convex set C in norm $\|.\|$ is defined as:

$$dist(z, C) = \inf\{\|z - x\| \quad | \quad x \in C\} \qquad (4)$$

We can find a unique projection vector in C to obtain a nonzero distance if $z \notin C$ or equal to zero if $z \in R^n$.

Definition 2: The closed convex hull (denoted as *convS*) of a nonempty set $S \subset R^n$ is the intersection of all closed convex sets containing S [16]:

$$convS = \{\theta_1 x_1 + \theta_2 x_2 + ... \theta_n, \sum_{i=1}^{n} \theta_i = 1\} \qquad (5)$$

The norm or distance used by the formulas above is usually represented as a linear product. In general, complex real-world applications require more expressive hypothesis spaces than a linear product. Kernel representations offer an alternative solution by projecting the data into a high dimensional feature space to increase the computational power of SVM. A kernel $K(x, z)$ and a feature map ϕ into a feature space F satisfying:

$$K(x, z) = <\phi(x), \phi(z)> \qquad (6)$$

For the linear separable problems, the kernel function can be expressed directly within the product of two vectors: $K(x, z) = x^t \times z$; for the linear inseparable problem, SVM adopts a non-linear kernel function (such as: *RBF* kernel) to map a linearly inseparable problem into a linearly separable one in Hilbert space. The distance between the two vectors in Hilbert space can be represented as follows [16]:

$$dis(x, z) = \|\phi(x) - \phi(z)\|^2$$
$$= <\phi(x) - \phi(z), \phi(x) - \phi(z)> \qquad (7)$$
$$= K(x, x) - 2K(x, z) + K(z, z)$$

From equations (4), (5) and (7), the distance of vector z to a convex hull *convS* can be represented as a projection:

$$dis \Pr ojection(z, convS) = \min_\theta \left\| z - \sum_{i=1}^{k} \theta_i x_i \right\|^2$$
$$= \min_\theta (K(z, z) - 2\sum_{i=1}^{k} \theta_i K(z, x_i) + \sum_{i=1}^{k}\sum_{j=1}^{k} \theta_i \theta_j K(x_i, x_j)) \qquad (8)$$
$$s.t. \sum_{i=2}^{k} \theta_i = 1, \; \theta_i \ge 0, \; i = 1, 2, ..., k$$

If we can find a group of θ_i to make the projection distance between z and *convS* equal to zero, it means the vector z is inside the *convS*. If the projection is not equal to zero, it means vector z is outside the *convS*. This formula can be used as an important criterion to construct the convex hull. Center of mass of the convex set in Hilbert space can be represented as [16]:

$$\phi(S) = \frac{1}{n}\sum_{i=1}^{n} \phi(x_i) \qquad (9)$$

The map $\phi(.)$ may be unknown in most of the kernel functions, so the center vector may cannot be obtained directly, but when the kernel function is given, the distance of a vector x to a set's center can be obtained, based on formulae (7) and (9):

$$disCenter(x,x) = \|\phi(x) - \phi(s)\|$$
$$= <\phi(x),\phi(x)> + <\phi(s),\phi(s)> - <\phi(x),\phi(s)> \quad (10)$$
$$= K(x,x) + \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}K(x_i,x_j) - \frac{2}{n}\sum_{i=1}^{n}K(x,x_i)$$

This formula can be used to evaluate the distance from a sample vector x to a set center in Hilbert space. Through formula (10) the distance of a vector x to a convex hull $convS$ can be represented as $disCenter(x,convS)$, which is a measure of importance to describe x's position in a convex hull.

B. Algorithms based on convex theory

Basing on convex theory and the distance formula we propose the following four algorithms:

(1) *Is_in_convex* algorithm: This algorithm tests whether or not a multi-dimensional vector x is in the convex decided by a sample set.

Is_in_convex (**Input:** a multi-dimensional vector *x*, sample set *S*)

Output: Boolean value (in or not in)

Begin

disProjection= disProjection (x,convS) by solving formula (8);

if *disProjection*<>0 { return true; }

else { return false; }

End

(2) *Get_convex_hull* algorithm: Given a sample set $S=\{x_1, x_2, \ldots, x_n\}$ with *n* multi-dimensional vectors, the convex hull *convS* decided by a set *S* can then be obtained.

Get_convex_hull (**Input**: A sample set *S*)

Output: convex hull *convS*

Begin

$$convS = convS \bigcup \{x_1, x_2\}; \quad S = S - \{x_1, x_2\};$$

foreach x_i in *S* {

$$dis_i = disCenter(x,convS)$$ by formula (10);

if (not ***Is_in_convex*** (x_k, convS)) { $$convS = convS \bigcup \{x_k\}; \}$$

}

return *convS*;

End

(3) *Sample_evaluation* algorithm: Given a multi-dimensional sample *x*, the training value (i.e. the important degree for SVM training) of sample *x* can be obtained, based on convex hull *convS*.

Sample_evaluation (**Input**: Sample *x*, convex hull *convS*)

Output: the training value

Begin

if (not (***Is_in_convex***(*x*, *convS*)) { return 1; }

Find x_{min} and its distance d_{min} by formula (7);

$$disCenter_x = disCenter(x,convS),$$

$$disCenter_{x\min} = disCenter(x_{\min},convS);$$

return $disCenter_x/(d_{min}+disCenter_{xmin})$;

End

(4) Based on above algorithms, the process of CTRSSSA is shown as follows

CTRSSSA (Input: Sample set *S*, Selection percentage *P*)

Output: Selected samples which have larger training values

Begin

convS= ***Get_convex_hull****(S)*;

$V =$ ***Sample_evaluation****(s,convS);*

num=number of samples in *S*;

snum=*num*P*;

orderedS= *S* rearranged with descending evaluation value (*V*) order;

return top *snum* samples in *orderedS*;

End

IV. EXPERIMENTS AND RESULTS

Our study area covers the whole Honghe National Nature Reserve (HNNR) which is located in the Sanjiang Plain (the largest fresh water wetland area in the northeast region of China). A Landsat-5 Thematic Mapper imagery acquired on August 31, 2009 (orbit number 113/27), sub image size 700×859) and six spectral bands of the image were downloaded, this image including blue (Band1), green (Band 2), red (Band3), near-infrared (Band4), and two mid-infrared (Band5 and 7). The composite image (bands 4, 3, 2) of whole image and the study area can be seen in Figure 2:
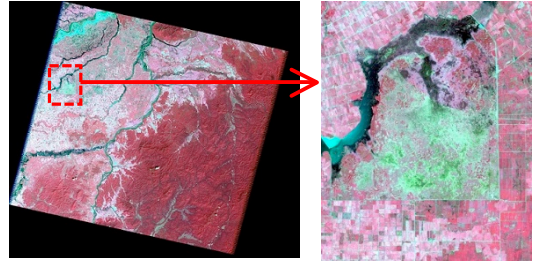

Figure 2. Remote Sensing Image and study area (composite of bands 4, 3, 2)

Based on field experience and investigation at study area, the study area's land cover/use categories include: Marsh Land (ML), Forestland (FL), Meadow (MD), Farmland (FD) and Water (WT), through Region Of Interest (ROI) tool, we select samples from study area image, samples with each category consisting of 1000 independent pixels, and 1000 samples of each category are further split into two sample sets: 200 samples as the training sample set and 800 samples as the testing sample set. The proposed algorithms are implemented in MATLAB R2011b, and LIBSVM 3.1 with its MATLAB interface adopted as the SVM classifier [17]. To evaluate the effectiveness of CTRSSSA for the sample selection, the proposed method is compared with the random sample selection (RSS) method.

In the experiment, selection percentage *P* varying from 100% to 1% in step 1% is adopted. Here, *P*=100%

signifies that all of the samples within the training set are selected, whereas *P*=1% denotes that only 10 samples in total (with 2 samples in each category) are chosen. SVM classifier model (a Linear Kernel) is trained by each group of the selected samples, and the classification accuracy is then tested with the corresponding testing set. The classification accuracies of the two sample selection methods (CTRSSSA and RSS) are shown in Table 1:

Table 1. Classification accuracy comparison between CTRSSSA method (the proposed method) and random sample selection (RSS) method

S%	C%	R%	S%	C%	R%	S%	C%	R%	S%	C%	R%
100	91.88	91.9	75	91.93	88.85	50	91.88	89.4	25	90.13	88.8
99	91.9	92.3	74	91.93	89.4	49	91.98	89.43	24	90.1	86.53
98	91.9	91.68	73	91.95	89.68	48	91.95	89.73	23	90.05	82.48
97	91.88	91.63	72	91.98	91.45	47	91.88	87.8	22	**90**	88.85
96	91.85	92.05	71	91.98	90.13	46	91.98	90.48	21	89.8	85.58
95	91.85	91.95	70	91.93	90.6	45	91.53	89.58	20	89.6	87.1
94	91.85	91.88	69	92	92.15	44	91.43	87.75	19	89.65	86.1
93	91.85	92.18	68	92.03	90.1	43	91.63	89.03	18	89.93	86
92	91.78	91.98	67	92.05	91.15	42	91.25	88.7	17	90.23	76.48
91	91.83	91.25	66	92.13	91.98	41	91.25	88.7	16	**90.18**	75.68
90	91.83	91.68	65	92.13	90.58	40	91.38	89.85	15	89.68	86.95
89	91.83	90.23	64	92.25	89.58	39	91.28	88.38	14	89.58	80.98
88	91.85	**89.75**	63	92.13	90.78	38	91.28	88.1	13	89.68	85.4
87	91.85	89.53	62	91.98	88.65	37	90.83	88.55	12	89.88	87.88
86	91.85	91	61	92.08	89.63	36	90.75	90.6	11	89.7	83.75
85	91.85	91.45	60	91.85	89.63	35	90.75	88.93	10	89.7	88.75
84	91.9	88.95	59	91.8	90.83	34	90.93	90.08	9	89.63	84.78
83	91.9	90.75	58	91.8	88.13	33	91	88.78	8	89.83	86.45
82	91.95	89.48	57	91.83	89.6	32	90.98	84.23	7	**87.95**	**70.7**
81	91.95	89.9	56	91.88	89.18	31	90.75	89.3	6	89.48	84.05
80	91.93	91.28	55	**91.88**	89.88	30	90.88	89.23	5	88.93	80.58
79	91.93	91.23	54	91.58	87.38	29	90.8	89.5	4	87.73	74.08
78	91.95	91.88	53	91.85	89.8	28	90.55	85.25	3	86.45	87
77	91.98	91.3	52	91.68	89.83	27	90.3	87.75	2	**86.18**	86.4
76	91.98	89.83	51	91.65	89.8	26	90.18	87.73	1	81.25	71.4

Nomenclature: selection percentage (S%), classification accuracy of CTRSSSA method (C%) and random sample selection method (R%).

The classification accuracy and comparison of two methods are shown in figure 3, 4 and 5:
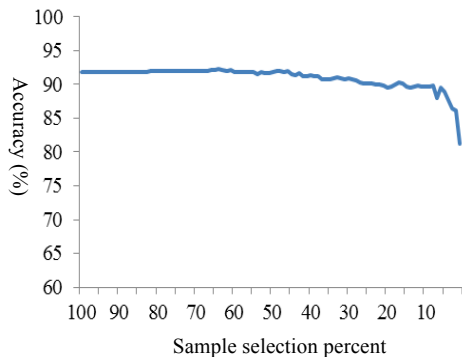

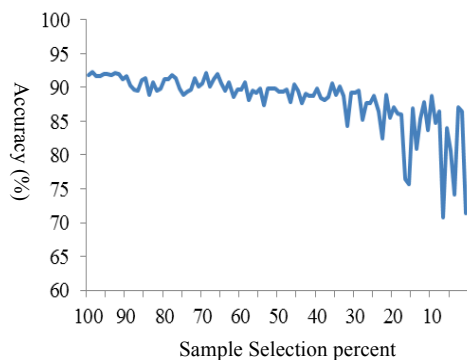Figure 3.   Classification accuracy of CTRSSSA method
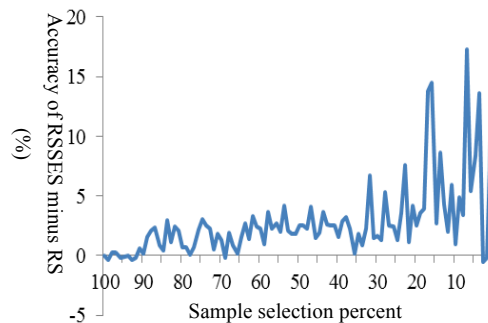

Figure 4.   Classification accuracy of RSS method

Figure 5.   Classification accuracy of CTRSSSA minus that of RSS

When the selection percentage *P*=100%, i.e. all of the training samples are selected, both CTRSSSA and RSS nearly attain the highest classification accuracy (92.3%). But along with the decrease in training samples, the classification accuracies of the two methods change in different patterns. For the CTRSSSA method, with a relatively stable and flat declining trend, the decline of the *P* of sample selection does not directly result in the decrease of classification accuracy (Fig 3). To be specific, compared with the classification accuracy of *P*=100%, an accuracy of 91.88% can still be achieved with a selection proportion *P*=55%; the classification is more than 90% accurate until *P*=22% (44 samples in each category); and 86.18% accuracy can still be reached until *P*=2% (just 4 samples in each category). The RSS method, however, a rapid classification accuracy decline is seen (the classification accuracy drops below 90% at P=88%). Moreover, the classification accuracy fluctuates remarkably and the smaller the selection percentage *P*, the more obvious the fluctuation (Fig 4). As we can see from

Fig 5 and Tab. 1, the CTRSSSA method can select more valuable training samples than RSS, and it achieves a higher classification accuracy in most cases (Fig 5), despite the fact that RSS's classification accuracy is slightly higher (only 0.55%) than CTRSSSA in P=100, 99, 96, 95, 94, 93, 92, 69, 3 and 2. Such an advantage becomes more and more notable with the decline of P, reaching the maximum (87.95%-70.7% = 17.25%) at P=7%.

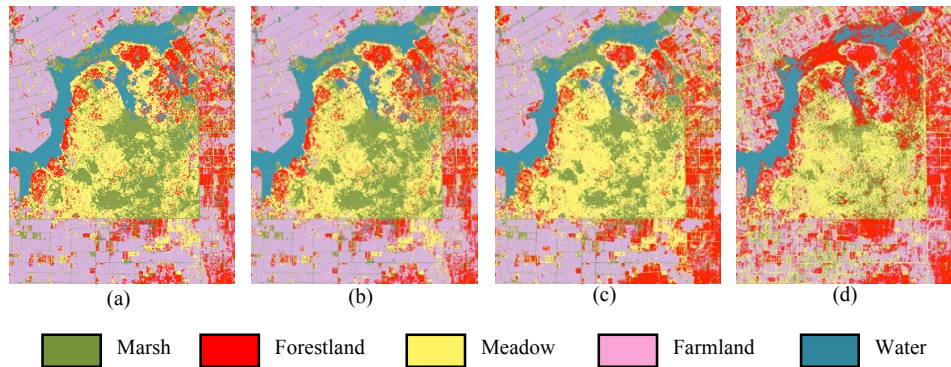| Marsh | Forestland | Meadow | Farmland | Water |

Figure 6.   Classification results: (a) 100% samples selected by CTBRSSSA; (b) 100% samples selected by RSS; (c) 7% samples selected by CTRSSSA; (d) 7% samples selected by RSS

Fig. 6 shows the classification results of the two methods for P=100% and P=7% respectively. When P=100%, both CTRSSSA and RSS select all samples and obtain approximately the best classification results, as illustrated by Fig 6.a (CTRSSSA) and Fig 6.b (RSS) in which almost all categories are correctly distinguished. However, when P=7% (5 categories with 14 samples in each, the training set consisting of 70 samples in total), RSS is clearly inferior to CTRSSSA. The classification accuracy of RSS is just 70.7%, with many categories being misclassified including Meadow, Farmland and even Water (Fig. 6.d), whereas the classification accuracy of CTRSSSA reaches 87%, with only partial Farmland being misclassified to Forestland and some Marsh to Meadow (Fig. 6.c).
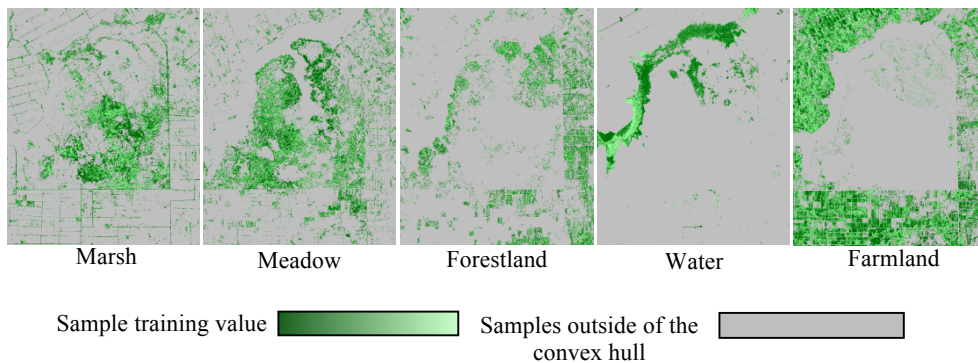

| Marsh | Meadow | Forestland | Water | Farmland |

Sample training value | Samples outside of the convex hull

Figure 7.   Evaluation results obtained by CTBRSSSA for each category

Fig. 7 shows the evaluation results and training value of each pixel on the remote sensing image by assigning the classified pixels with green or gray color. Therein, the dark green color represents the pixels' position in the corresponding category's convex hull, and the corresponding color depth from dark to light reflects the magnitude of training value (i.e., the darker color, the larger the training value).The gray color represents the pixels outside of the corresponding convex hull, CTBRSSSA tend to select darker green pixels, which are easily misclassified in small sample sizes. Selecting darker green pixels can bring SVM classifier more accuracy when fewer samples are selected.

V.   CONCLUSION

SVM is a widely used remote sensing image classifier that is data-dependent, and the quality of its training sample set greatly influences the classification result.

In this paper, convex theory is first introduced into the selection process of the remote sensing training sample to quantitatively describe the relationship or importance between a sample and a convex hull quantitatively. Three algorithms, namely, *Is_in_convex*, *Get_convex_hull*, and *Sample_evaluation*, are designed. *Is_in_convex* tests whether or not a multi-dimensional vector x is in the convex. *Get_convex_hul*l can obtain the convex hull from a sample set. *Sample_evaluation* algorithm can calculate a sample's training value in the interval [0, 1]. A larger

value corresponds to a more important training sample, where 1 refers to the most important sample and 0 the least important sample. With the help of the samples' training value, CTRSSSA enables the selection of the most valuable samples for the SVM classifier, and it can maintain high classification accuracy, even when fewer samples are selected and utilized.

Our experiments, where a group of samples from 100% to 1% are selected, demonstrate that in most cases, samples that are more valuable can be selected and higher classification accuracy can be achieved by CTRSSSA compared with RSS method.

CTRSSSA is more stable than RSS, with a slower declining trend in classification accuracy along with the decrease in sample selection percentage. This statement is still true even when the number of samples is rather small. Furthermore, the fluctuation trend of CTRSSSA is less severe than that of RSS. With the help of CTRSSSA, users can select fewer and more valuable samples when classifying a remote sensing image and increase the SVM training speed to obtain better classification results.

VI. ACKNOWLEDGMENT

REFERENCES

[1] G. Mountrakis, J. Im, C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol.66, no.3, pp. 247-259, May 2011. http://dx.doi.org/10.1016/j.isprsjprs.2010.11.001

[2] Z. Zhou, X. Wang, G. Pan, "CMCCR: Classification Based on Multiple Class-Correlation Rules," *Journal of Digital Information Management*, vol.10, no.2, pp. 64-70, April 2012.

[3] L. Su, "Optimizing support vector machine learning for semi-arid vegetation mapping by using clustering analysis," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol.64, no.3, pp. 407-413, July 2009. http://dx.doi.org/10.1016/j.isprsjprs.2009.02.002

[4] A. Lyhyaoui, M. Martinez, I. Mora, M. Vazquez, J. Sancho, A. Figueiras-Vidal, "Sample selection via clustering to construct support vector-like classifiers," *IEEE Transactions on Neural Networks*, vol. 10, no.6, pp. 1474-1481, August 1999. http://dx.doi.org/10.1109/72.809092

[5] A.M. Brisha, "A Comparison Between Methods of Selecting Cluster Head," *International Journal of Online Engineering*, vol.6,

no.4, pp.7-13, November 2010. http://dx.doi.org/10.3991/ijoe.v6i4.1325

[6] S. Abe and T. Inoue, "Fast training of support vector machines by extracting boundary data," *Artificial Neural Networks-Icann 2001 Proceedings* ,vol 2130, no.1, pp. 308-313, August 2001.

[7] A. Adamatzky, "On excitable β-skeletons," *Journal of Computational Science,* vol.1, no.3, pp.175-186, August 2010. http://dx.doi.org/10.1016/j.jocs.2010.07.003

[8] J. Wang, P. Neskovic, L. Cooper, "Training data selection for support vector machines," *Advances in Natural Computation*, vol. 7626, no.1 , pp. 554-564, August 2005. http://dx.doi.org/10.1007/11539087_71

[9] G. Foody and A. Mathur, "Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification," *Remote Sensing of Environment*, vol. 93, no.1 , pp. 107-117, 2004. http://dx.doi.org/10.1016/j.rse.2004.06.017

[10] G. Foody, A. Mathur, C. Sanchez-Hernandez, D. Boyd, "Training set size requirements for the classification of a specific class," *Remote Sensing of Environment* ,vol. 104, no.1, pp. 107-117, September 2006. http://dx.doi.org/10.1016/j.rse.2006.03.004

[11] A. Singla, S. Patra, L. Bruzzone,"A novel classification technique based on progressive transductive SVM learning,"*Pattern Recognition Letters*, vol.42, no.1, pp.101-106, June 2014. http://dx.doi.org/10.1016/j.patrec.2014.02.003

[12] T. Temuçin, H. Tozan, J. Valíček, M.Harničárová, "A fuzzy based decision support model for non-traditional machining process select," *Technical Gazette*, vol.20,no.5, pp 787-793, October 2013.

[13] S. Shima, M. Todoriki, A. Suzuki, "SVM-based feature selection of latent semantic features,"*Pattern Recognition Letters*, vol.25, no.9, pp. 1051-1057, July 2004. http://dx.doi.org/10.1016/j.patrec.2004.03.002

[14] S. Theodoridis and M. Mavroforakis, "Reduced convex hulls: A geometric approach to support vector machines," *IEEE Signal Processing Magazine*, vol.24, no.1, pp. 119-122, May 2007. http://dx.doi.org/10.1109/MSP.2007.361610

[15] P. Bertsekas, *Convex Optimization Theory*, Massachusetts, Athena Scientific Press, pp. 21-30, June 2009.

[16] S. Boyd, and L. Vandenberghe, *Convex Optimization*, United Kingdom, Cambridge University Press, pp.40-45, March 2004.

[17] R. Fan, P. Chen, C. Lin, . "Working set selection using second order information for training SVM," *Journal of Machine Learning Research,* Vol.6, no.3, pp. 1889-1918, December 2005.

AUTHORS

Xin Pan(Corresponding author) is an associate professor in School of Computer & Information Technology, Changchun Institute of Technology, Changchun ,130012, China (email: 101103991@qq.com)

Hongbin Sun is a professor in School of Computer & Information Technology, Changchun Institute of Technology, Changchun, 130012, China (email: 53866661@qq.com)