DETECTING HUMAN ACTIVITIES IN THE ARCTIC OCEAN BY CONSTRUCTING AND ANALYZING SUPER-RESOLUTION IMAGES FROM MODIS DATA

Shizhi Chen and YingLi Tian
Department of Electrical Engineering
The City College of New York
New York NY, USA
{schen21, ytian}@ccny.cuny.edu

William S. Weiss
Naval Undersea Warfare Center,
1176 Howell St., Newport, RI 02841.
william.s.weiss@navy.mil

ABSTRACT

In this paper, we propose a framework to detect human activities by constructing super-resolution images from the MODIS data. The highest resolution of the MODIS images is 250 meters per pixel, which is usually not enough to detect human activities. By magnifying and de-blurring the low resolution MODIS image through the Support-Vector Regression, the constructed super-resolution image can achieve 4 to 8 times higher resolution than the original MODIS image. To evaluate the feasibility of the super-resolution MODIS images for the application of human activity detection, we collect a dataset by selecting four land cover types through Google Earth: the land with human activities, the land without human activities, the water without ice, and the land covered with snow and ice. Using a learning-based method, surface reflectance from the super-resolution MODIS image predicts land cover type of a geo-location specified by latitude and longitude. Experimental results demonstrate feasibility of the proposed approach for human activity detection using the super-resolution MODIS images.

KEYWORDS: Satellite Image, Human Activity Detection, Arctic Ocean, Super-Resolution, MODIS Data

INTRODUCTION

Detecting and locating human activities can play a very important role in the surveillance of the Arctic Ocean and can be particularly critical to future Navy operations

Researchers in the remote sensing community have tried to directly use NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) satellite data to detect land cover changes caused by human activities [4, 8, 10].

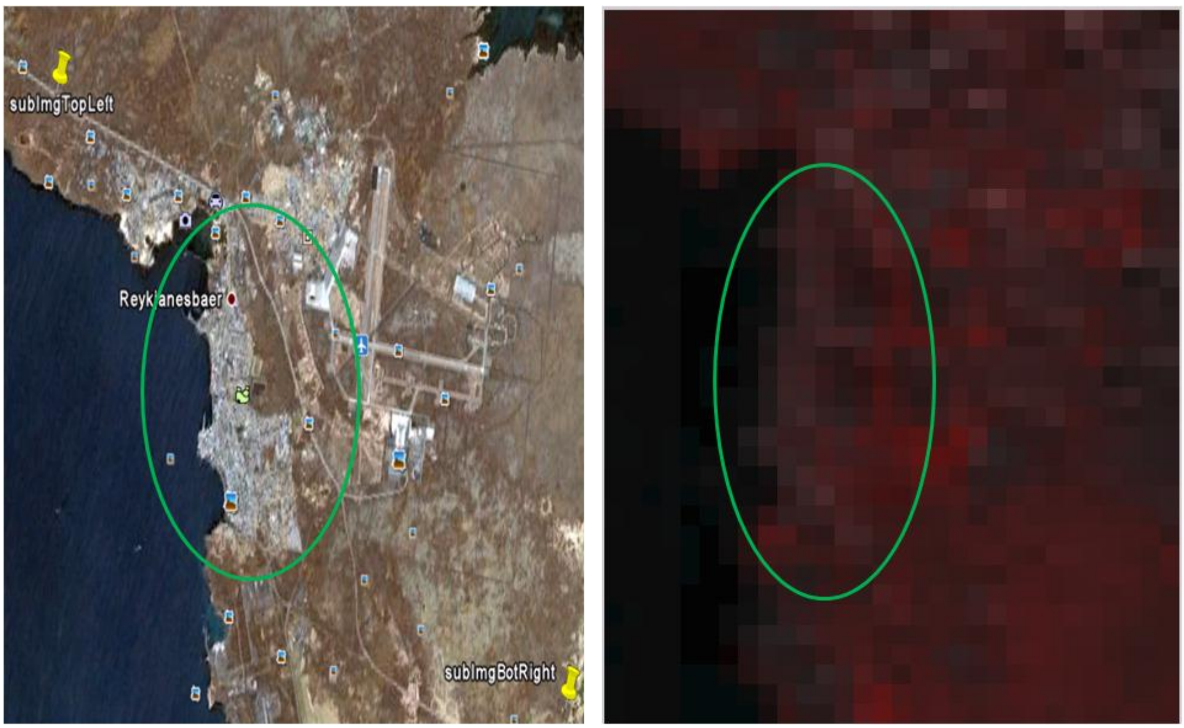
Zhan *et al*. [10] applied the Vegetative Cover Conversion (VCC) algorithms based on MODIS 250m radiance data to detect land cover changes caused by human activities or extreme natural events. Specifically, they used the MODIS images to detect wild fires, deforestation and floods, *etc*. Their experiments show that MODIS imaging is effective when significantly large portion of land cover is changed.

Pilloni *et al*. [8] detected land cover changes between 2005 and 2008 in Sardinia by utilizing a 16-day composite of the MODIS 250m data. Their studies suggest that the MODIS data is useful to evaluate land cover changes in heterogeneous vegetation areas, which can potentially apply to the monitoring and planning of environmental management.

However, the size of most human activities in the Arctic Ocean, such as surfaced submarines, ice camp structures, ice breaker trail widths, oil platforms, *etc*., is often less than 100 meters. Hence, the highest resolution of the MODIS data at 250 meters can potentially limit the applications of the MODIS data for human activity detection. Figure 1(a) shows a region with the human activities in Reykjavik, Iceland from Google Earth. Figure 1(b) shows the corresponding MODIS image, which is the false color RGB combination of band 2 (infrared), band 1(red), and band 1. The corresponding region with the human activities is circled in green.

Super-resolution has been an active research area in the computer vision community. Its goal is to estimate a high resolution image from one, or a set, of low resolution images. Two types of popular approaches are interpolation-based [6, 9] and learning-based super-resolution [1, 5, 7].

Li and Orchard [6] propose an edge-directed interpolation algorithm to construct a high resolution image from a low resolution image. They first estimate local covariance coefficients for the low resolution image. Then these covariance estimates are adapted to the interpolation of the high resolution image according to the geometric duality between the low resolution covariance and the high resolution covariance.



(a) (b)

Figure 1: (a) Human activities in Reykjavik, Iceland from Google Earth; (b) The corresponding MODIS image, which is the false color RGB combination of band 2 (infrared), band 1(red), and band 1.

Liu *et al*. [7] synthesize a high resolution face image from a low resolution face image by developing an algorithm which satisfies three constraints, *i.e.*, sanity constraint, global constraint, and local constraint. The reconstructed high-resolution face image shows high similarity to the original high resolution face image.

Inspired by these advances on super-resolution technology, we propose a Support-Vector regression based super-resolution algorithm, *i.e.*, SRSVR. The SRSVR only uses a single low resolution image as the input and generates a higher resolution image by learning a de-blurring process when down-sampling. Then we apply the SRSVR to synthesize the high resolution MODIS image from its original low resolution version, *i.e*., 250m per pixel. The constructed super-resolution image can achieve 4 to 8 times higher resolution, which can potentially avoid the resolution limitation on human activity detection in the Arctic Ocean.

To evaluate the possibility of applying the super-resolution MODIS images to human activity detection, we collect a database, which consists of four land cover types, *i.e*., (1) the land with human activities, (2) the land without human activities, (3) the water without ice, and (4) the land covered with snow and ice. The ground truth of land cover type at each selected geo-location is labeled by observing Google Earth, and verified in the MODIS images that these selected locations are not occluded by clouds.

By employing the Support-Vector Machine (SVM) learning method, the surface reflectance at the red and the infrared bands from the super-resolution MODIS image predict land cover type for a geo-location specified by latitude and longitude. Experimental results demonstrate feasibility of the proposed approach for human activity detection using the super-resolution MODIS images.

METHOD OF DETECTING HUMAN ACTIVITIES FROM MODIS DATA

Method Overview

Figure 2 shows a flowchart to evaluate the feasibility of human activity detection using the SRSVR super-resolution algorithm. We collect a database consisting of four land cover types, *i.e.*, (1) the land with human activities (Human), (2) the land without human activities (Land), (3) the water without ice (Water), and (4) the land covered with snow and ice (Ice).

We select 20 sample geo-locations in the area of Iceland for each land cover type through Google Earth. Then we verify if these locations in the MODIS images are clear of clouds. Local features, *i.e.*, neighboring pixels' surface reflectance, are extracted from the super-resolution images, which are synthesized through the SRSVR algorithm from the original 250m resolution MODIS images. Finally, we employ the SVM classifier to recognize land cover type, *i.e.*, "Human", "Land", "Water" and "Ice".

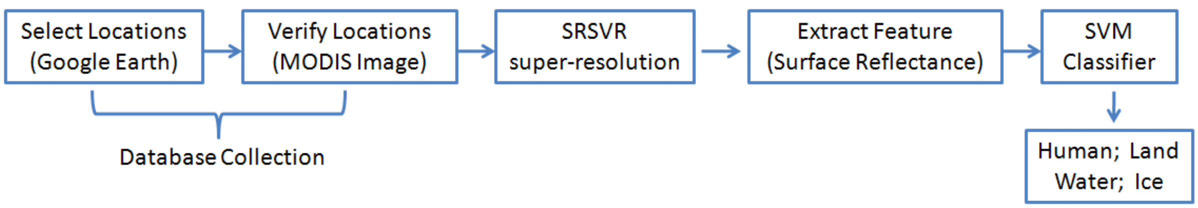
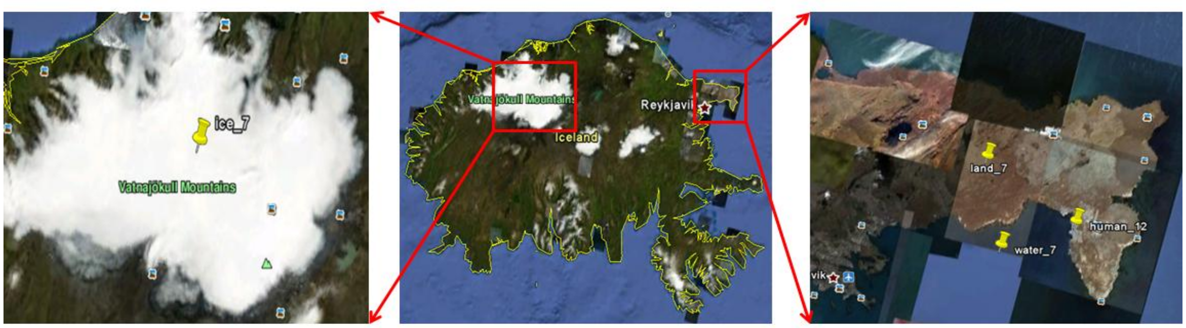
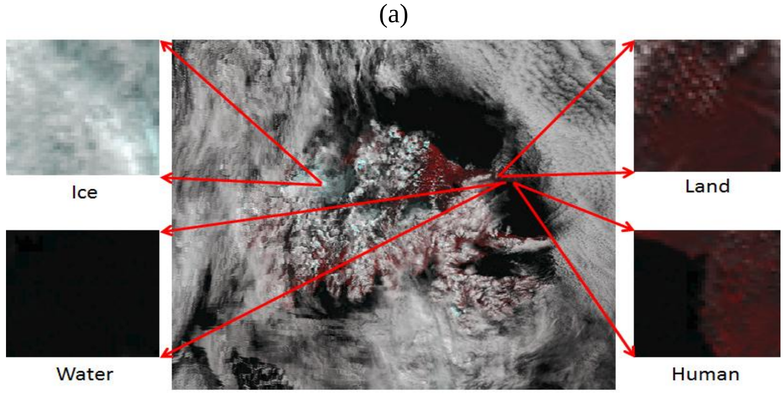


Figure 2: Flowchart to evaluate the feasibility of human activities detections using the SRSVR super-resolution algorithm.



(a)



(b)

Figure 3: (a) Sample geo-locations of "Human", "Land", "Water", and "Ice" from Google Earth; (b) The false color MODIS image at the corresponding locations.

Database Collection

For each of the four land cover types, we sample 20 geo-locations in the area of Iceland through Google Earth. Figure 3 (a) shows sample geo-locations of “Human”, “Land”, “Water”, and “Ice” from Google Earth. Then we select three MODIS images captured by the satellites Terra and Aqua, which are clear of clouds at the selected geo-locations, specified by latitudes and longitudes. The MODIS images contain surface reflectance at band 1 (red) and band 2 (infrared) with 250m resolution. Figure 3(b) shows the false color MODIS image at the corresponding geo-locations.

Super-Resolution Image Generation Using Support Vector Regression

We develop a super-resolution algorithm based on the Support Vector regression, *i.e.*, SRSVR, to magnify and de-blur the 250m resolution MODIS images. The SRSVR algorithm follows the super-resolution algorithm proposed in [5], except that we adopt the Support Vector regression (SVR) instead of the Gaussian process regression. The SRSVR algorithm takes a single low resolution image and generates a de-blurred high resolution image. The constructed super-resolution image can achieve 4 to 8 times higher resolution as compared to the original 250m resolution. Hence, super-resolution images can provide more details on human activity detection in the Arctic Ocean.

The SRSVR algorithm consists of two major steps: (1) up-sampling and (2) de-blurring as shown in Figure 4(a) and 4(b) respectively. In the up-sampling step, we perform bi-cubic interpolation on the input low resolution image (LR) with a desired scaling factor. Then we partition both low resolution image and the interpolated high resolution image into corresponding overlapped patches. For each pair of the corresponding patches, we sample pixels in the low resolution patch as training targets with eight neighbors of each sampled pixel as their training feature vectors. After performing the Support-Vector Regression (SVR) on the obtained targets and the feature vectors, we obtain the SVR model. The SVR model is then used to predict the blurred high resolution (HR) image using the eight neighbors of the bi-cubically interpolated high resolution image as its input feature vector.

In the de-blurring step, we further blur and down-sample the blurred HR image to obtain the blurred LR image. Similar to the up-sampling step, the blurred HR image, the blurred LR image, and the original LR image are partitioned into overlapped patches as shown in Figure 4(b). For each patch of the original LR image, we sample pixels as the training targets, and the eight neighbor pixels in the corresponding blurred LR patch as the feature vector for each sampled pixel. We perform the SVR regression and obtain the SVR model, which has modeled the de-blurring process at low resolution. The SVR model is then used to predict the de-blurred HR image using the eight neighbor pixels of the blurred HR image.

We magnify the MODIS images by applying the SRSVR super-resolution on the surface reflectance at band 1 (red) and band 2 (infrared) respectively.

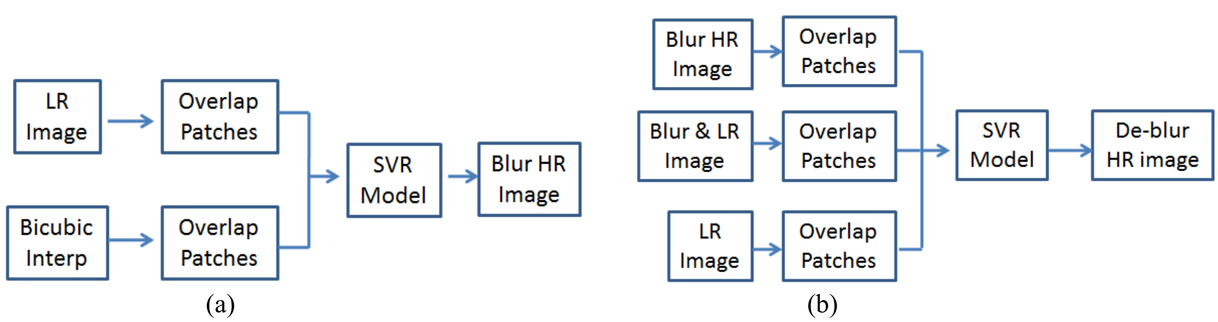


Figure 4: Flowchart of the SRSVR super-resolution algorithm in two steps: (a) up-sampling and (b) de-blurring.

Feature Extraction

For each selected pixel, we extract surface reflectance of a patch in the super-resolution MODIS image at band 1 and band 2 respectively. The patch width is the scaling factor multiplied by 3. The pixel intensity in the patch is sensitive to rotation change and noise data. Hence we propose to use a statistic based feature representation. That is to extract the mean and the variance of the super-resolution patch at band 1 and band 2 respectively.

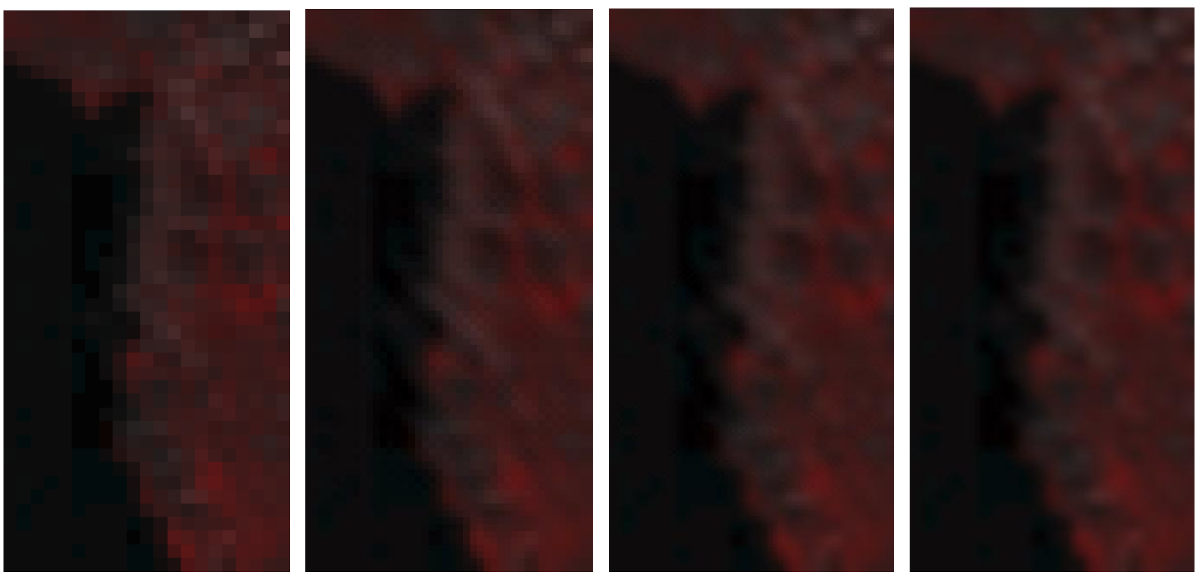
Support Vector Machine based Land Cover Type Classification

The Support Vector Machine (SVM) is used to find an optimal hyper-plane to separate a binary class data with maximum margin [2, 3]. We employ the RBF kernel to project the feature vector into the high dimensional space. The RBF kernel usually achieves state of the art performance in object recognition literature. The maximum voting strategy is used to solve our multi-class classification problem. That is to classify “Human”, “Land”, “Water” and “Ice” land cover types according to the surface reflectance at band 1 and band 2 of the MODIS data.

EXPERIMENTS

Results of SRSVR Super-Resolution Image Reconstruction

Figure 5(a) shows the original 250m resolution false color MODIS image with band 2, band 1, and band 1 as its RGB components. Figure 5(b)-(d) show its high resolution MODIS images with the scaling factor of 2, 4 and 8 respectively. The super-resolution images obtained from the SRSVR algorithm show more details on edges or coast lines, where human activities often occur. However, as the scaling factor increases to 8, the super-resolution MODIS image is not showing continuous improvement from the visual inspection.



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

Figure 5: Super-resolution through the SRSVR algorithm: (a) original 250m resolution; (b) super-resolution by the scale of 2; (c) super-resolution by the scale of 4; (d) super-resolution by the scale of 8.

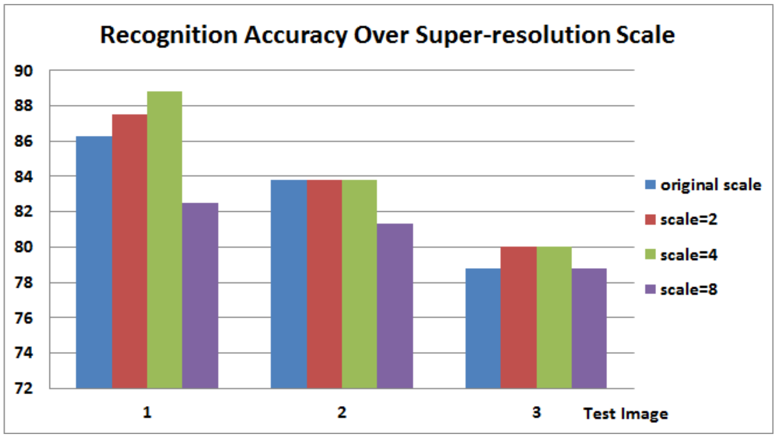
Results of Human Activity Detection

In order to evaluate the feasibility of the SRSVR super-resolution algorithm for human activity detection, we conduct experiments using the collected database, which consists of four land cover types, i.e., (1) the land with human activities (Human), (2) the land without human activities (Land), (3) the water without ice (Water), and (4) the land covered with snow and ice (Ice). Out of the three selected MODIS images, we choose one as testing data and the remaining two as training data. Three-fold cross validation is performed in the experiment.

Figure 6(a) shows the comparison of the classification accuracy over different scaling factors. The average accuracy improvement at the scale of 4 is more than 1.2% higher than the original 250m resolution. This result suggests that super-resolution can indeed improve the human activity detections. As the scale increases to 8, the recognition accuracy starts decreasing.

The confusion table in Figure 6(b) shows the details of the classification accuracy. As expected, “Water” and “Ice” land cover types are both classified with high accuracy. The land with Human activities (Human) and the land without human activities (Land) are confused with each other. However, our approach can still achieve 80% true positive rate for the land cover type with human activities. The average accuracy over the four land cover types is approximately 84% at the super-resolution scale of 4.

As mentioned in last section, we use the mean and the variance of the super-resolution patch as the feature representation. The statistic based feature representation is invariant to rotation change and noise data in the MODIS images. Figure 7 shows the comparison between the statistics based feature representation and the intensity based feature representation, *i.e.*, the feature vector formed by the pixel intensities in the super-resolution patch. The statistics based feature representation shows significantly better performance over all three testing MODIS images.



(a)

	human	land	water	ice
human	16	4	0	0
land	5	15	0	0
water	0	0	20	0
ice	0	0	0	20

(b)

Figure 6:.(a) Average classification accuracy of the four land cover types over different super-resolution scale; (b) confusion matrix at the scaling factor of 4.

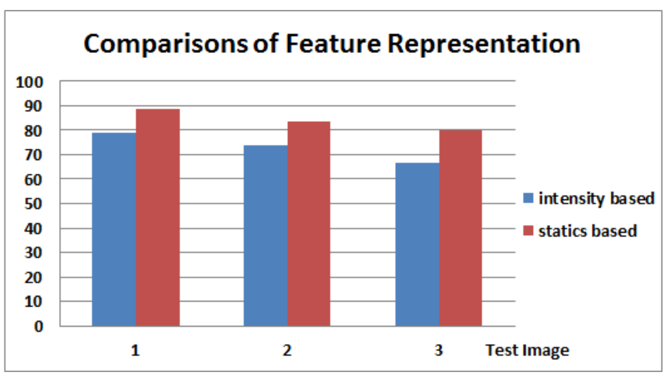


Figure 7: Comparison between the statistics based feature representation with the intensity based feature representation at the super-resolution scale of 4.

CONCLUSION

In this paper, we have proposed a super-resolution algorithm using the Support-Vector regression, *i.e.*, the SRSVR. We have studied the feasibility of applying the SRSVR super-resolution algorithm to human activities detection in the Arctic Ocean. We collected and annotated a database of four land cover types including land with human activities. The experimental results on the database demonstrate that the super-resolution MODIS images can achieve a true positive rate of 80% on human activity detection. It outperforms the original 250m resolution MODIS data by 1.2% in accuracy when up-sampling the MODIS images by 4 times. The results suggest the feasibility of the super-resolution algorithm, the SRSVR, on the application of human activity detection in the Arctic Ocean.

ACKNOWLEDGMENTS

This manuscript was supported in part under the ONR Summer Faculty Program and ARO grant W911NF-09-1-0565. Shizhi Chen is funded by NOAA Cooperative Remote Sensing Science and Technology (CREST) Fellowship.

References:

1. S. Baker and T. Kanade, “Limits on super-resolution and how to break them”, IEEE Transaction on Pattern Analysis and Machine Intelligence, 2002.
2. C. Burges, “A tutorial on support vector machines for pattern recognition”, Data Mining and Knowledge Discovery, 2(2), 121–167, 1998.
3. C. Chang and C. Lin, LIBSVM: a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.
4. R. Fries, M. Hansen, J. Townshend, R. Sohlberg, “Global land cover classifications at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers”, International Journal Remote Sensing, 1998.
5. H. He and W. Siu, “Single image super-resolution using Gaussian process regression”, Computer Vision and Pattern Recognition, 2011.
6. X. Li, M. Orchard, “New edge-directed interpolation”, IEEE Transactions on Image Processing, 2001.
7. C. Liu, H. Shum, and C. Zhang, “A two-step approach to hallucinating faces: global parametric model and local nonparametric model”, Computer Vision and Pattern Recognition, 2001.
8. M. Pilloni, M. Melis, and A. Marini, “Analysis and validation of a methodology to evaluate land cover change in the mediterranean basin using multitemporal MODIS data”, Present Environment and Sustainable Development, 2010.
9. W. Tam, C. Kok, and W. Siu, “Modified edge-directed interpolation for images”, Journal of Electronic Imaging, 2010.
10. X. Zhan, R. Sohlberg, J. Townshend, C. DiMiceli, M. Carroll, J. Eastman, M. Hansen, and R. DeFries, “Detection of land cover changes using MODIS 250m data”, Remote Sensing of Environment 83, 336-350, 2000.