# Mars image segmentation with most relevant features among wavelet and color features

Abdolreza Rashno
Department of Electrical and Computer Engineering, Isfahan University of Technology,
a.rashno@ec.iut.ac.ir

Mohamad Saraee
School of Computing, Science and Engineering, University of Salford-Manchester, United Kingdom,
m.saraee@salford.ac.uk

Saeed Sadri
Department of Electrical and Computer Engineering, Isfahan University of Technology,
sadri@cc.iut.ac.ir

***Abstract*—Mars rover is a robot which explores the Mars surface, is equipped to front-line Panoramic Camera (Pancam). Automatic processing and segmentation of images taken by Pancam is one of the most important and most significant tasks of Mars rover since the transformation cost of images from Mars to earth is extremely high. In this paper, a new feature vector for image pixels will be proposed as well as a new feature selection schema based on ant colony optimization(ACO). Then, the most relevant features are presented for multiclass Support Vector Machine (SVM) classifier which led to high accuracy pixel classification and then image segmentation. Our proposed method is compared with genetic algorithm feature selection, experimental results show that the proposed method outperforms this method in the terms of accuracy and efficiently.**

***Keywords—Pixel classification, image segmentation,feature selection, ant colony optimization, genetic algorithm, support vector machine;***

## I. INTRODUCTION

Recently the automated robot, named as Mars Rover is configured in Mars environment,is equipped with Pancam instrument which could be assumed as scientist's eyes on Mars. Automated image analysis system with high degree of accuracy for analyzing of Mars images is required to escape from manually analyzing which is time-consuming task with human errors[1].

Our contribution in this research is automatic detection and segmentation of different types of rocks and sands from their surrounding background in Mars images. The first step in implementation of such systems is the extraction of feature set from image pixels. Despite the fact that which features are the best descriptor for pixels classes and lead to more accurate classification results, there are so many features proposed by researches as being a candidate in feature set[2,3]. Although the more relevant features in feature set, the higher accuracy in detection and segmentation could be achieved, increasing dimension of feature set involve high computational time which is not in our interest especially in on-board applications such as Mars rover with the limitation of power, memory and processing rate. On the other hand, there is the idea that not all extracted features are relevant in classification [4]. A key issue is that the size of the feature vector should be kept as small as possible since most of the classification methods used for Mars rover involve high computational costs when feature vectors move towards higher dimensionalities. Feature selection is a method that removes redundant and irrelevant feature components. As a result, low time complexity and high system accuracy could be achieved. Many feature selection approaches have been proposed for different applications such as our proposed speaker verification schema [5], face recognition [6], data mining and pattern recognition[7] and so on. Feature selection has also been applied to Mars rover automated system in the term of minimizing feature subset by Fuzzy-rough feature selection[1,8] and information gain ranking technique[9] in mars image detection and segmentation.

In this paper, the new schema for rock detection and segmentation is proposed. Firstly, the feature set including, wavelet features, color descriptor features, color statistic features and color local histogram features are all extracted from image. Then, the new feature selection method based on ant colony optimization will be applied to extracted feature set and the most relevant and most significant features are selected from whole feature set. Finally, the multiclass support vector machine classifier is adapted for this application and trained by the way that classifier parameters are estimated with highest classification accuracy. The segmented Mars images are the result of our proposed system.

The rest of this paper is organized as follows: Section 2 describes the extraction of feature set from images. Pixel classifier and image database are explained in more details in Section 3. Feature selection and our proposed method are described completely in Section 4. Experimental setup and results are described in Sections 5 and 6 respectively. Finally, the conclusion and future works are discussed in Section 7.

## II. FEATURE EXTRACTION

### *A. Wavelet Features*

Wavelet transform could be applied to images as 2-dimensional signals in Eq. (1).

$$CWT(s,a,b) = \frac{1}{\sqrt{s}} \iint f(x,y)\, \Psi\left(\frac{x-a}{s}, \frac{y-b}{s}\right) dxdy \qquad (1)$$

The frequency components of the image are obtained up to k level. In each level, these components are LL that is approximation of image and HL, LH and HH that are

978-1-4799-8733-7/15/$31.00 ©2015 IEEE

horizontal, vertical and diagonal frequency details respectively [10].

In this research the low and high frequency components are used for extracting features of window around pixels since some type of rocks such as layered and wave-shaped rocks with having vertical and horizontal frequency components are easily segmented by these types of features. We use Frobenius norm of frequency components which for a vector $M = [M_1, M_2, \dots, M_n]$ is defined in Eq. (2).

$$M \in R^n \implies \|M\|_F = (\sum_{i=1}^{n}(M_i)^2)^{\frac{1}{2}} \qquad (2)$$

For Mars images in RGB space, first they are converted to gray space, then, first level of Haar filter for LH and HL components are computed. Then, Frobenius norms of rows of LH, columns of LH, rows of HL and columns of HL are computed and denoted by $a_i, b_i, c_i$ and $d_i$ respectively.Finally, the mean and standard deviation of vectors $a_i, b_i, c_i$ and $d_i$ are computed for wavelet features of each pixel. This method aims for reducing sensitivity toward local differences existing in each class of pixels. These differences can result in increasing distances between different classes and decreasing them within a same class. Extracting features in this way improves the classification accuracy.

## B. Color Features

### 1) Dominant color descriptor

Dominant color descriptor(DCD) is one of the approved color descriptors in the MPEG-7 Final Committee Draft among several number of histogram descriptors. Both representative colors and the percentage of each color are included in DCD. Moreover,DCD provide an effective and compact color representation, could be applied for color distribution in an image or a region of interesting [11]. In DCD, first, each color is divided into the number of partitions named as course partitions. Then, all points in a same partition are assumed to be similar and near to each other. Partition centers are the average value of all pixels in each partition and are calculated with the Eq. (3).

$$C_j = \frac{\sum_{p \in P_i} p}{\sum_{p \in P_i} 1} \qquad (3)$$

Which $P_i$ is an ith partition. In this research the DCD features are extracted in RGB domain. Each pixel color value is replaced by the center value of partition which is belonged to it. As a result, applying DCD quantize images which the possible colors for each pixel is equal with the number of partitions. For a window around a pixel, the dominant colors for R, G and B components are computed and represented as DCD features for that pixel.

### 2) Local color histogram features

Histogram is a bar showing the number of pixels falling in grey levels named resolution ranges or bins[12]. Since Mars images are RGB with the range 0-255 in each component, 256 bins is needed per R, G and B components which is computationally expensive. This range divided to 8 equal sub-ranges for each component. Since our interest is local features means that features of window around a pixel, 8 features are extracted for each component of a region of image. So, each pixel of image is mapped to 24-dimensional feature vector in the term of local color histogram features.

### 3) Color statistic features

First and second moments of window around pixels are mean and standard deviation respectively. We have applied these features in RGB domain. For R, G and B the mean and standard deviation are computed as color statistic features.

# III. Pixel Classifier

## A. Support Vector Machine

SVM is one of the powerful machine learning methods based on statistical learning theory. SVM is widely used in pattern recognition problems because of its good generalization ability compared with traditional classification methods [13]. SVM was originally proposed for binary classification problems but it can also be generalized for multi-class applications. The problem is to find a hyper-plane to separate instances from two classes. If the input data are not linearly separable, we can't find this hyper-plane. Thus, input data are mapped into higher dimensional space using map functions in order to create linear separability between instances[14].

## B. Image Database

This research proposes an image classifier which classifies and segments the different type of rocks of Panorama images from Mars environment. The image with 180-degree of scan named as Home Plate South panorama image [15] is used in this research. This image is taken by the panoramic camera on NASA's Mars Exploration Rover Spirit, acquired on October 2007, has the dimension of 23123x3775 pixels with about 90 meters length. A part of this image are show in Fig.1.

# IV. Feature Selection

Feature selection is a discrete optimization problem which selects *m* features among *n* ones that $m \leq n$ [16]. Alleviating the effect of the dimensionality, enhancing generalization capability, improving model interpretability and speeding up the learning process are all the benefits of removing irrelevant and redundant features from the main features of data [5 ,17]. Some popular evolutionary-based feature selection approaches are ant colony optimization (ACO)[18], particle swarm optimization (PSO)[19], genetic algorithm (GA) [20] and Relief [21].

Fig.1.A part of Home Plate South panorama image.

### A. *Genetic Algorithm for Feature Selection*

GA introduced by Holland in 1975, is a randomized heuristic search technique based on biological evolution strategies. GA has been applied to our problem of interest in several approaches as well as its application in feature selection and feature weighting. The purpose of GA in feature selection task is finding an optimal binary vector with the smallest number of 1s such that the classifier performance is maximized[20 , 22]. Another application of GA is feature weighting which assigns numerical weights to features instead of binary 0 and 1 weights[23].

### B. *Ant Colony Optimizationfor Feature Selection*

As a solution of hard optimization problems, an iterative and probabilistic meta-heuristic method named ACO was proposed by Dorigo and colleagues which simulates the natural behavior of ants, consisting of mechanisms of adaptation and cooperation [24]. Similar to GA, ACO can be reformulated for feature selection problem by the way that finds a path with minimum cost in graph. In contrast with classic ACO, here, nodes in the graph represent features and the edges between nodes denote the choice of the next feature [25]. The pheromone update rule and transition rule of ACO algorithm can be used by a little reformulation for feature selection. In this case, pheromone and heuristic value are associated to nodes(features) rather than edges. The probability that ant *k* selects feature *i* at time step *t* is calculated by Eq. (4).

$$P_i^k(t) = \begin{cases} \frac{|\tau_i(t)|^\gamma |\eta_i|^\delta}{\sum_{u\in J^k}|\tau_u(t)|^\gamma |\eta_u|^\delta} & if\ i \in J^k \\ 0 & otherwise \end{cases} \quad (4)$$

Where $J^k$ is the set of features that are allowed to be added to the partial solution if they are not visited so far.$\tau_i(t)$and$\eta_i$ are the pheromone value and heuristic desirability associated with feature *i* respectively. $\gamma$and$\delta$ are weights of pheromone value and the heuristic information respectively[26]. The amount of deposited pheromone of ant *k* on feature *i* in step *t* is calculated by Eq. (5)[17].

$$\Delta\tau_i{}^k(t) = \begin{cases} \phi.\mathrm{H}(\mathrm{S}^k(t)) + \frac{\psi.(n-|S^k(t)|)}{n} & \text{if } i \in S^k(t) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Where *n* is the number of all features, $S^k(t)$ and $|S^k(t)|$ are the feature subset and its length found by ant $k$ at iteration *t* respectively. $H(S^k(t))$is the evaluation of subset $S^k(t)$ which is the classifier performance in this literature, $\phi$and $\psi$are parameters that determine the importance of classifier performance and feature subset length respectively. Finally, the pheromone of nodes are updated by the Eq. (6)[26].

$$\tau_i(t+1) = (1-\rho).\tau_i(t) + \sum_{k=1}^{m}\Delta\tau_i{}^k(t) + \Delta\tau_i{}^g(t) \quad (6)$$

Where $m$ is the number of ants,$\rho$is an evaporation rate which is constant and *g* is the best ant in previous iteration. It means that node pheromones are affected by all ants and more affected by the best ant which deposits additional pheromone on nodes. This causes the search of ants to stay around the optimal solution in the next iterations[17].

### C. *Proposed feature selection algorithm based on ACO*

In the proposed method, feature selection is applied to all pixel classes simultaneously. This mean that the feature selection schema try to select most relevant features by the way that maximize the classification accuracy of all pixel classes simultaneously. So, an optimal feature vector is presented for all pixel classes.

The following algorithm describes the proposed feature selection schema in more details.

1. Initialization:

- Set population size for ants.
- Assign each feature to one node in graph

and set random intensity of pheromone for it.

- Set the maximum number of iteration.

2. Dividing the features into groups based on table 1.

3. Assign a feature to each ant randomly and mark this feature as “visited feature”.

4. For an ant, select one feature from each feature group randomly and save them in array *a*. From features in *a*, select the feature with highest probability which is calculated by (4) and mark this feature as ‘visited’.

5. Check the ant for reaching to its threshold, if not reached, go to 4.

6. For all ants, do steps 3-4.

7. Each ant deposits a quantity of pheromone on features which marked as ‘visited’ in its path by Eq. (7).

$$\Delta\text{pheromone}(i,k) = \propto.\left(\text{Model Accurarcy}(k)\right) + \beta.\left(\frac{\text{FeaturesNumber}-\text{FeaturesNumber}(k)}{\text{FeaturesNumber}}\right) \quad (7)$$

Where $\Delta\text{pheromone}(i,k)$ is an amount of pheromone deposit by ant $k$ on feature $i$, model accuracy is the F-measure of model trained by features founded by ant $k$, FeatutesNumber and FeaturesNumber(k) are the number of all features and the number of features in the path of ant $k$ respectively. Finally, $\alpha$ and $\beta$ are two parameters that control the relative weight of classifier performance and feature subset length respectively that $\alpha + \beta = 1$.

8. Find the best ant, an ant with the highest F-measure.

9.Pheromone update is calculated in Eq. (8).The pheromone of all features is evaporated, all ants deposit a pheromone on the features in their paths and the best ant has extra effect on pheromone update.

$$\text{pheromone}_i(t+1) = (1-\rho).\text{pheromone}_i(t) + \sum_{k=1}^{m}\Delta\text{pheromone}(i,k) + \Delta\text{pheromone}(i,\text{BestAnt}) \quad (8)$$

10. Previous ants are removed and new ants are generated randomly.

11. If the stop criteria is achieved (all ants follow the same path) or maximum number of iteration is reached go to 12 else go to 3

12. The ant with maximum F-measure has the best solution.

13. End

In this task the stop criteria for ants is a threshold which defined by Eq. (9).

$$\text{Ant\_Threshold} = \varphi * \exp^{-\frac{FN}{N}} + \omega * \exp^{\text{F\_Measure}} \quad (9)$$

Where *FN* is the feature cardinality of the selected feature by the ant so far, *N* is the number of all features, $\varphi$ and $\omega$are the parameters that control the effect of feature size and F_measure respectively and $\varphi + \omega = 1$. Afterwards, all ants complete their search as a term of finding feature subsets. The pheromone update is then started by 3 following rules: each ant deposit some pheromone on features, the idea which stem from the natural behavior of ants, some pheromone is evaporated by time and finally, the best ant update its path pheromone stronger since it could be a sign for other ants to follow its path in next steps. The amount of pheromones in all features in ant path is affected by the F_measure and the number of features selected by that ant.

It is clear that ACO feature selection is a wrapper method among two types of feature selection as filter and wrapper. In these types of feature selection methods, accuracy of classification is involved in feature selection task. It means that the model must be trained and tested for feature subsets in each step of feature selection. Each ant after randomly selection of initial feature, select next feature based on the efficiency of that feature. Suppose that feature vector has *n* features. Each ant must train the model n-1 times for selection of second feature. After that, n-2 times of model training is needed for the selection of third feature. This continues until the ant exit from its traverse. In the worst case, each ant train the model $\frac{n(n-1)}{2}$ times in each iteration. The computation cost is explosively high. Although the feature selection is done in the offline phase, once for classification, we propose the new procedure for decreasing this cost with preserving the accuracy of classification. After many experiments it was cleared that it is no need that each ant,at each iteration, evaluate all non-visited features for select one of them which lead to model training and its high computation cost. To escape from this issue we divided features into groups as shown in table 1.

The idea of grouping features has this benefit that each ant for selection of each feature in each step of it's traverse just select among 13 features (one feature taken from each group randomly) rather than complete feature set. It is important since selection of each feature involve model training and testing. This idea stem from the fact that features in each group have similar behavior for classification, so, it is not need for evaluation all features in each step. On the other hand, all features of a group have a chance for being selected by ants. For example, if LH_meanL is the selected feature from group 1 at step *k* for a ant, feature LH_stdL will be a candidate for being selected by that ant from steps k+1 onwards. This shows that all features have a chance of being selected by ants.

## V. EXPERIMENTAL SETUP

To show the utility of the proposed feature selection algorithms a series of experiments is conducted. We implement our algorithms on a machine with 2.26 GHz Corei7 CPU and 6GB of RAM and windows 7. The Home Plate South panorama image with the dimension of 23123x3775 pixels with about 90 meters length is divided to 512x512 sub images and then 7 different classes for pixels are defined for the task of pixel classification and image segmentation. Firstly the sub images are zero padded by 10 pixels per side. Each pixel is windowed with the 21x21 window around it, for the window of each pixel, features are computed and each pixel is mapped to its corresponding feature vector. For a window around each pixel, wavelet decomposition is computed for gray image, norm of rows and columns of LH and norm of rows and columns of HL lead to 4 vectors. The mean and standard deviation of these 4 vectors are 8 wavelet features per pixel. In color features, the number of color dominant is set to 8, so, 24 features are computed for R, G and B components. Also, mean and standard deviation of R, G and B components are 6 color statistic features. Finally, with 4 bins for the histogram of R, G and B, 12 features are computed. So, for each pixel, a feature vector of dimension 50 is used for pixel classification.

For class *i* number of pixels from this class are selected as positive samples and then the *k* randomly selected samples used as samples of negative class. The SVM is trained for these positive and negative classes which we named the SVM model for pixel class *i*. Table 2 shows the number of positive and negative samples for all pixel classes used in this research.

TABLE I. GROUPING FEATURES OF A PIXEL IN ACO FEATURE SELECTION

| Group Name | Group Features | Group Name | Group Features |
|---|---|---|---|
| Group 1 | **LH_meanL and LH_stdL** | Group 8 | $R_m$ *and* $R_s$ |
| Group 2 | **LH_meanH and LH_stdH** | Group 9 | $G_m$ *and* $G_s$ |
| Group 3 | **HL_meanH and HL_stdH** | Group 10 | $B_m$ *and* $B_s$ |
| Group 4 | **HL_meanLnadHL_stdL** | Group 11 | $R_i$**: All histogram bins of R component** |
| Group 5 | $\mathbf{DCDR}_i$ **: All color dominants of R component** | Group 12 | $G_i$**: All histogram bins of G component** |
| Group 6 | $\mathbf{DCDR}_i$ **: All color dominants of G component** | Group 13 | $B_i$**: All histogram bins of B component** |
| Group 7 | $\mathbf{DCDR}_i$ **: All color dominants of B component** | | |

TABLE II. NUMBER OF POSITIVE AND NEGATIVE SAMPLES FOR ALL PIXEL CLASSES

| | Class 1 | Class 2 | Class 3 | Class 4 | Class 5 | Class 6 | Class 7 |
|---|---|---|---|---|---|---|---|
| Number of positive samples | 5651 | 5470 | 4941 | 10771 | 8682 | 4959 | 5475 |
| Number of negative samples | 5734 | 4365 | 5412 | 12544 | 9322 | 4566 | 7165 |
| Sum | 11385 | 9835 | 10353 | 23515 | 18004 | 9525 | 11640 |

TABLE III. PARAMETER SETTINGS FOR GA AND THE PROPOSED ALGORITHMS

| Methods | Iteration | Population | Initial pheromone | Crossover Probability | Mutation Probability | $\delta$ | $\gamma$ | $\alpha$ | $\beta$ | $\rho$ | $\varphi$ | $\omega$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GA | 100 | 100 | - | 0.6 | 0.008 | - | - | - | - | - | - | |
| Proposed Algorithm | 100 | 100 | 1 | - | - | 1 | 1 | 0.6 | 0.4 | 0.3 | 0.1 | 0.9 |

In these experiments, various parameter values were tested for GA and the proposed algorithms. According to our experiments, the highest performance in each method is achieved by setting the parameters to values shown in Table 3.

For classification task, the learning method is set to sequential minimization optimization(SMO) with maximum iterationnumber of 20000 and kernel cache limit to 1000. The multilayer Perceptron(MLP) with scale [0.01,-0.01] and linear functions are used as kernel functions in this research since these kernels had the best accuracies for classification.

## VI. EXPERIMENTAL RESULTS

Our proposed pixel classification and image segmentation system is evaluated by precision, recall and F-measure as the harmonic mean of precision and recall.

Our proposed model is compared with genetic algorithm(GA) which is an evolutionary algorithm. Dropping irrelevant features and presenting most relevant ones for SVM models corresponding to pixel classes, surprisingly increase the classification accuracy a little. Also, results show that our proposed algorithm has better results than genetic algorithm for two SVM kernels and different pixel classes. The precision, recall and F-measure of our proposed method is compared with genetic algorithm, are shown in table 4.

The final results of this research is pixel classification and image segmentation. All 7 types of rocks[1] including Rover tracks, small black stone and sand, medium black stone and sand, layered rocks, wave rocks, dark large size rocks with shadow and flat rocks are all segmented in Home Plate South sub images. However, some parts of images which were far away fromPancom camera are not segmented to one of mentioned 7 rock types in this research since their details are not visible. As it is clear from segmented images, the main error source of segmentation is stem from boundary between regions. Since the feature vector of a pixel is features of window around that pixel, the pixels which are near to region boundaries are windowed by pixels from other classes, so they incorrectly classified to other class labels. However, our proposed feature selection methods, select the features which are not depended to boundary as possible, this increases the accuracy of segmentation even after dimension reduction of pixel features. We assigned 8 colors including red, green, blue, yellow, indigo, pink, gray and violet for classes 1-7 and non-segmented parts respectively. Fig. 2(a-c) shows some samples of segmented images.

TABLE IV. CLASSIFICATION ACCURACY FOR SECOND PROPOSED METHOD AND GA

| | | Complete Feature set | | Selected Features by proposed method | | Selected Features by GA | |
|---|---|---|---|---|---|---|---|
| | | SVM_MLP | SVM_Linear | SVM_MLP | SVM_Linear | SVM_MLP | SVM_Linear |
| **C1** | **Precision** | 0.6139 | 0.6553 | 0.5889 | 0.6670 | 0.6122 | 0.6542 |
| | **Recall** | 0.9874 | 0.9511 | 0.9892 | 0.9407 | 0.9619 | 0.9483 |
| | **F-measure** | **0.7571** | **0.7760** | **0.7390** | **0.7805** | **0.7482** | **0.7743** |
| | **Precision** | 0.9975 | 0.9991 | 0.9960 | 0.9990 | 0.9935 | 0.9632 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **C2** | **Recall** | 0.8600 | 0.8348 | 0.8883 | 0.8380 | 0.8415 | 0.8143 |
| | **F-measure** | **0.9237** | **0.9096** | **0.9391** | **0.9114** | **0.9112** | **0.8825** |
| **C3** | **Precision** | 0.9952 | 0.9981 | 0.9953 | 0.9927 | 0.9713 | 0.9854 |
| | **Recall** | 0.7520 | 0.6850 | 0.7779 | 0.7096 | 0.7455 | 0.6834 |
| | **F-measure** | **0.8567** | **0.8124** | **0.8732** | **0.8276** | **0.8436** | **0.8071** |
| **C4** | **Precision** | 0.9062 | 0.9486 | 0.9077 | 0.9513 | 0.8965 | 0.9589 |
| | **Recall** | 0.9440 | 0.9174 | 0.9517 | 0.9175 | 0.9576 | 0.9200 |
| | **F-measure** | **0.9247** | **0.9323** | **0.9293** | **0.9341** | **0.9260** | **0.9390** |
| **C5** | **Precision** | 0.8705 | 0.8524 | 0.8945 | 0.8437 | 0.8633 | 0.8586 |
| | **Recall** | 0.9357 | 0.9423 | 0.9317 | 0.9454 | 0.9042 | 0.9199 |
| | **F-measure** | **0.9020** | **0.8951** | **0.9127** | **0.8916** | **0.8833** | **0.8882** |
| **C6** | **Precision** | 0.8080 | 0.9373 | 0.8037 | 0.9373 | 0.7854 | 0.9025 |
| | **Recall** | 0.9020 | 0.5824 | 0.8992 | 0.5745 | 0.8945 | 0.5474 |
| | **F-measure** | **0.8524** | **0.7184** | **0.8487** | **0.7124** | **0.8364** | **0.6815** |
| **C7** | **Precision** | 0.9926 | 0.9400 | 0.9863 | 0.9037 | 0.9932 | 0.9165 |
| | **Recall** | 0.8187 | 0.8537 | 0.8247 | 0.8713 | 0.7913 | 0.8465 |
| | **F-measure** | **0.8973** | **0.8948** | **0.8983** | **0.8872** | **0.8808** | **0.8801** |
| **Average Precision** | | **0.8835** | **0.9036** | **0.8817** | **0.8992** | **0.8736** | **0.8913** |
| **Average Recall** | | **0.8856** | **0.8231** | **0.8946** | **0.8281** | **0.8709** | **0.8114** |
| **Average F-measure** | | **0.8734** | **0.8475** | **0.8771** | **0.8492** | **0.8613** | **0.8361** |
| **Optimal feature subset found by proposed feature selection method and SVM classifier** | | | | LH_meanL, HL_stdH, HL_stdL, DCDR1,DCDR5,DCDG4,DCDB1,Bm,Bs,R1,R6, G2,G3, B1,B4,B6,B7 | | | |

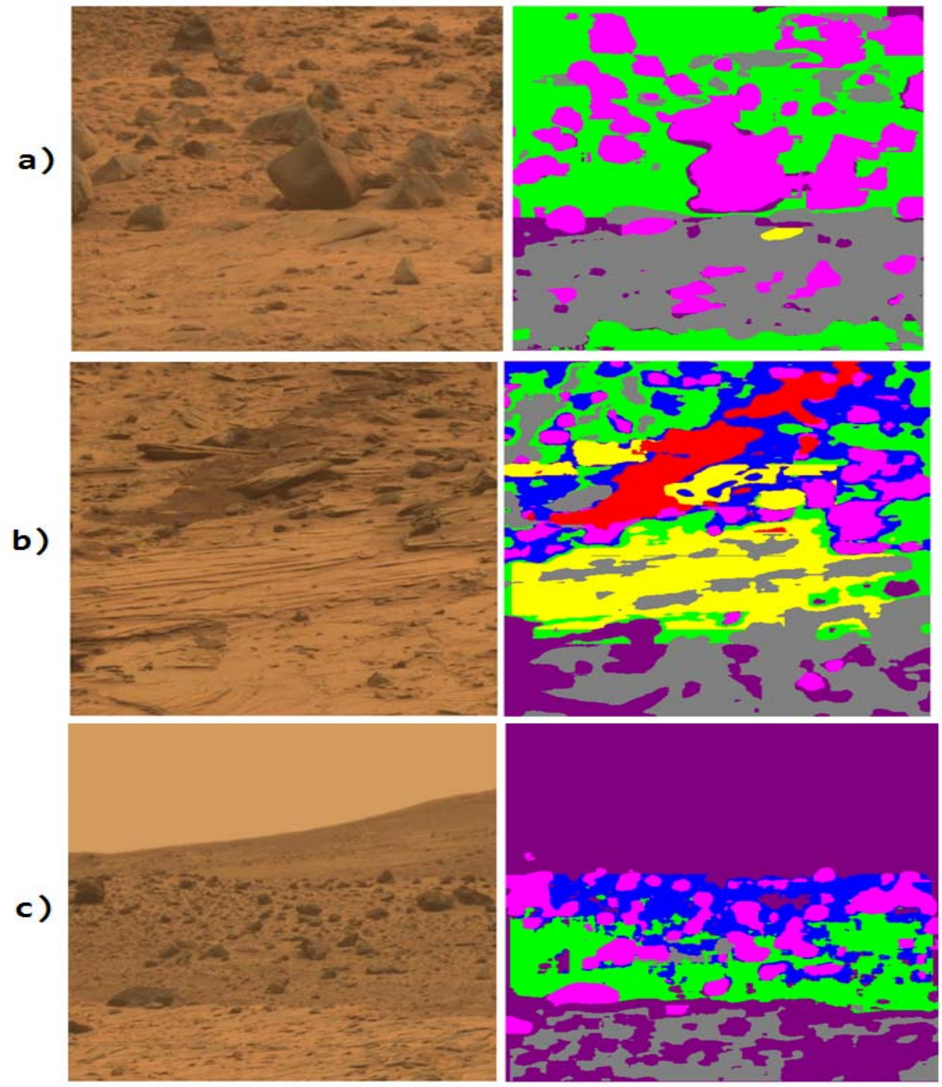


Fig.2. Some samples of segmented images

## VII. CONCLUSION AND FUTURE WORKS

Home plate south panorama images contain information about Mars, taken by mars rover, a robot which explore on Mars surface. In this paper, a new pixel classification schema was proposed which leads to Mars image segmentation. Firstly, each pixel is windowed and mapped to a feature vector include wavelet and color features. The type of extracted features was proposed by the way that all pixel classes could be discriminated from other classes and appropriate segmentation of images are achieved. Then, most relevant features among complete feature set were selected by our proposed feature selection schema based on ant colony optimization which worked with classification performance of all pixel classes. Finally, optimal feature sets were presented for multi-class SVM based pixel classification schema. Classification of pixels segments the images into pixel groups. Our proposed model for image segmentation was compared with genetic algorithm; results showed that the accuracy of our method is better than genetic algorithm feature selection. For future works, the performance of the proposed approach can be evaluated by taking into account other classifiers. Other feature selection methods can be improved and applied to such systems. In addition, intrinsic property of data such as relieff weights can be used in population-based techniques such as ACO, GA and Particle swarm optimization (PSO) algorithms for faster convergence. Finally, other features like shape features, texture features and color features in other color domains could be applied for pixel classification task.

## REFERENCES


[1] Shang, Changjing, and Dave Barnes. "Fuzzy-rough feature selection aided support vector machines for mars image classification." *Computer Vision and Image Understanding* 117.3 (2013): 202-213.

[2] K. Huang, S. Aviyente, Wavelet feature selection for image classification, IEEE Transactions on Image Processing 17 (9) (2008) 1709–1720.

[3] Kachanubal, Tossaporn, and Somkait Udomhunsakul. "Rock textures classification based on textural and spectral features." *Int J Comput Intell* 4 (2008): 240-246.

[4] D. Puig, M. Garcia, Automatic texture feature selection for image pixel classification, Pattern Recognition 39 (11) (2006) 1996–2009.

[5] Rashno Abdolreza, Hossein SadeghianNejad, and Abed Heshmati. "Highly Efficient Dimension Reduction for Text-Independent Speaker Verification Based on Relieff Algorithm and Support Vector Machines." *International Journal of Signal Processing, Image Processing & Pattern Recognition* 6.1 (2013).

[6] Gundimada, S., Asari, V. K., & Gudur, N. (2010). Face recognition inmulti-sensor images based on a novel modular feature selection technique. Information Fusion, 11, 124-132

[7] Jensen, R. (2005). Combining rough and fuzzy sets for feature selection. Ph.D. thesis, University of Edinburgh.

[8] R. Jensen, Q. Shen, New approaches to fuzzy-rough feature selection, IEEE Transactions on Fuzzy Systems 17 (4) (2009) 824–838.

[9] C. Shang, D. Barnes, Combining support vector machines and information gain ranking for classification of mars mcmurdo panorama images, in: Proceedings of the 17th IEEE International Conference on Image Processing, 2010, pp. 1061–1064.

[10] Yildizer, Ela, et al. "Integrating wavelets with clustering and indexing for effective content-based image retrieval." *Knowledge-Based Systems* 31 (2012): 55-66.

[11] ISO/IEC 15938-3/FDIS Information Technology—Multimedia Content Description Interface—Part 3 Visual Jul. 2001, ISO/IEC/JTC1/SC29/WG11 Doc. N4358.

[12] R. Gonzalez, R.E. Woods, Digital Image Processing and Scene Analysis, third ed., Prentice Hall, 2008.

[13] Vapnik, V. (1995). The Nature of Statistical Learning Theory. (2nd ed.). Springer-Verlag, London, UK, (Chapter 5).

[14] Burges, C. J. C. (1999). Geometry and Invariance in Kernel Based Method, Advance in Kernel Method-Support Vector Learning. MIT Press, Cambridge, MA, pp, 86–116.

[15]http://marswatch.astro.cornell.edu/pancam_instrument/ home_plate_south.html.

[16] Mladenic , D. (2006). Feature selection for dimensionality reduction. Subspace, latent structure and feature selection, statistical and optimization,perspectives workshop. Lecture Notes in Computer Science, Springer, 3940, 84–102.

[17] Nemati, S., & Basiri, M. (2011). Text-independent speaker verification using ant colony optimization-based selected features. Expert Systems with Applications, 38, 620–630.

[18] Tabakhi, Sina, Parham Moradi, and Fardin Akhlaghian. "An unsupervised feature selection algorithm based on ant colony optimization." *Engineering Applications of Artificial Intelligence* 32 (2014): 112-123.

[19] Wang, X., Yang, J., Teng, X., Xia, W., & Jensen, R. (2007). Feature selection based on rough sets and particle swarm optimization. Pattern Recognition Letters, 28(4), 459–471.

[20] Tsai, Chih-Fong, William Eberle, and Chi-Yuan Chu. "Genetic algorithms in feature and instance selection." *Knowledge-Based Systems* 39 (2013): 240-247.

[21] Kira, K., & Rendell, L. A. (1992). A Practical Approach to Feature Selection. The 9thInternational Conference on Machine Learning,pages 249–256.

[22] Yang, J., & Honavar, V. (1998). Feature subset selection using a genetic algorithm. IEEE Intelligent Systems, 13, 44–49.

[23] Beddoe, Gareth R., and Sanja Petrovic. "Selecting and weighting features using a genetic algorithm in a case-based reasoning approach to personnel rostering." *European Journal of Operational Research* 175.2 (2006): 649-671.

[24] Dorigo, M., & Caro, G. D. (1999). Ant colony optimization: A new meta-heuristic. In Proceedings of the congress on evolutionary computing.

[25] Lung Huang, C. (2009). ACO-based hybrid classification system with feature subset selection and model parameters optimization. Neurocomputing, 73, 438–448.

[26] RashidyKanan, H., & Faez, K. (2008). An improved feature selection method based on ant colony optimization (ACO) evaluated on face recognition system. Applied Mathematics and Computation, 205, 716–725.