

A selective combined classification algorithm for mapping alterations on ASTER data

Majid M. Oskouei · Wolfgang Busch

Received: 1 March 2011 / Accepted: 5 January 2012 / Published online: 26 January 2012

© Società Italiana di Fotogrammetria e Topografia (SIFET) 2012

Abstract We present a method for combined classification aiming to map alterations in a set of ASTER (Advanced Spaceborne Thermal Emission and reflection) data in the Erongo Complex, Namibia. Ten alterations detected by the matched filtering unmixing method on the Hyperion dataset of the area are therefore used as training classes. The separability of the classes was computed to evaluate the ability of ASTER data to spectrally discriminate between these classes. The outcome of this computation is satisfactory for the high-probability training dataset. In order to improve the accuracy of upcoming processes, classes with high similarity (low separability) were combined. The classification of ASTER scene is then performed with the use of both individual and combined classification classifiers. A new combined classification method (named selective combined classification (SCC)) was developed in this research to achieve the highest possible accuracy in the resultant classification map. An accuracy analysis has proven the advantages and capability of SCC among all classifiers tested in this study (both individual and combined).

Keywords Remote sensing · ASTER · Classification · Alteration mapping · Selective combined classification

Introduction

Supervised classification approaches have frequently been used for mapping alterations and rock types in the geological studies, for instance, by Kruse et al. (2002), Kruse

(2002), Rowan and Mars (2003), Paskaleva et al. (2004), Favretto and Geletti (2004), Hewson et al. (2005), Galvao et al. (2005), Vaughan et al. (2005), Hubbard and Crowley (2005), as well as Wang and Zhang (2006). The reasons for selecting a particular classification method, however, are rarely discussed. In this study, therefore, we performed an accuracy analysis to make a comparison between different approaches. The confusion matrix is then formed, and the accuracy of the results of the classifiers is computed based on the location of training classes.

In addition to individual classifiers, combined classification systems are sometimes used for improving the overall classification process. The theoretical and experimental results reported in the predicable studies, however, have clearly emphasized that combined classifiers are effective only if the individual classifiers are accurate and diverse, that is, if they exhibit low error rates and make different errors. Previous studies have indicated that the creation of accurate and diverse classifiers is a very difficult task (Sharkey et al. 2000; Giacinto and Roli 2001). In this paper, a new approach is developed by taking accuracy differences of base classifiers into account. This is designated as the selective combined classification (SCC), since this method makes use of selective application of rule images from base classifiers.

Study area

The area under investigation is located in the northwest of Namibia and includes the Erongo Complex with a diameter of approximately 35 km, which is one of the largest Cretaceous anorogenic complexes in that country. The centre of the complex is located approximately at 21°40'S and 15°38'E. Figure 1 illustrates band 3 of the Advanced Spaceborne Thermal Emission and Reflection (ASTER) image of the complex.

This complex represents the eroded core of a caldera structure with peripheral and central granitic intrusions. Surrounding the outer granitic intrusions of the Erongo Complex is a ring dyke of olivine dolerite, which locally attains some 200 m

M. M. Oskouei (✉)
Mining Engineering Faculty, Sahand University of Technology,
Tabriz, Iran
e-mail: moskouie@yahoo.com

W. Busch
Department of Mine Surveying and Geomatics,
Clausthal University of Technology,
Clausthal-Zellerfeld, Germany

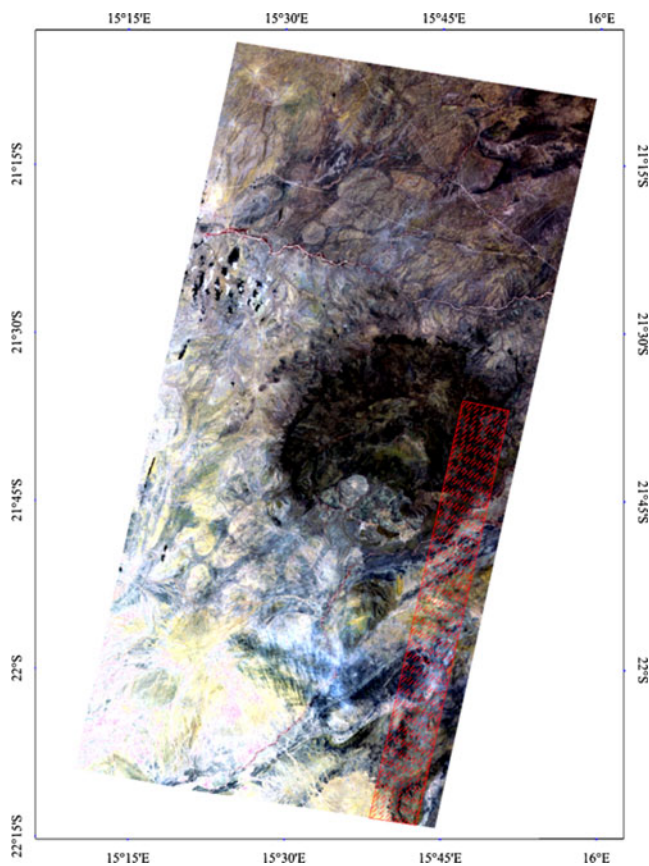


Fig. 1 ASTER colour composite (3, 2, 1) scene of Erongo complex and overlapping area of Hyperion image on it

in thickness and has a radius of 32 km. The ring dyke weathers easily and is therefore, highly eroded. The central part of the Erongo Complex consists of a layered sequence of volcanic rocks, which form prominent cliffs rising several hundred metres above the surrounding basement. The interior of the complex is deeply eroded and thus gives access to the roots of

the structure. The basement rocks consist of mica schists and meta-greywackes of the Kuiseb Formation and various intrusions of granites. In the southeast, the rocks of the Erongo Complex overlie the Triassic Lions Head Formation, which consists of conglomerates, gritstone, arkose with interbedded siltstone and mudstone, as well as quartz arenite (Schneider 2004).

Data description

ASTER datasets of Erongo consist of two northern and southern scenes. ASTER data generally includes 14 channels in three VNIR, SWIR, and TIR wavelength ranges. Ground resolutions of those three groups are different (VNIR, 15 m; SWIR, 30 m; TIR, 90 m), therefore, two datasets were produced by resampling and stacking all channels. The resultant datasets are two images, including 14 channels in respect to ASTER wavelengths and 15-m ground resolution. Two northern and southern scenes mosaiced to take place a solid dataset (Fig. 1). The data was then atmospherically corrected (for VNIR and SWIR bands) using FLAASH algorithm and is corrected for crosstalk error as well. The normalized difference vegetation index (NDVI) of the dataset was determined using by channels 2 and 3. This is important to mask pixels with high NDVI to improve the classification outcomes. Therefore, the pixels with more than 0.18 NDVI were masked before classification.

Training classes

The resultant map of alteration zones produced by matched filtering unmixing method on Hyperion dataset is used as training pixels for the classification task in this study. Table 1 illustrates detected minerals for each distinguished end-

Table 1 Detected minerals for each end-member using spectral feature fitting and their alteration type

End-members	Minerals	Alteration
1	Calcite _{2,1,3} –siderite–prochlorite _{2,1} –corrensite–chlorite	Carbonate
2	Alunite _{2,6,4,5} –rivadavite–ammonio illite–ulexite _{1,2}	Argillic (illite–alunite)
3	Antigorite,brucite	Magnesium skarn
4	Ammonio illite	Argillic (illite)
5	Mascagnite _{2,1} –sphalerite _{4,5} –acmite–alunite _{2,6,4} –pectolite ₁ –buddingtonite _{2,1} –ammonio chlorite–fluorapatite–ulexite ₂ –augite ₁	Calcic skarn
6	Alunite _{2,4,6,5} –rivadavite	Argillic (alunite)
7	Mascagnite ₂ –buddingtonite _{2,1} –ammonio jarosite	Sodic
8	Ulexite _{1,2} –alunite _{2,6,4,5,1} –rivadavite–ammonio illite	Argillic
9	Kainite–natrolite _{1,2} –syngenite–goethit ₁	Zeolitic
10	Alunite _{4,2,6} –rivadavite–orthoclase ₃ –ulexite _{1,2} –ammonio illite	Adularia

Table 2 Separability computations for maximum map classes

Classes	2	3	4	5	6	7	8	9	10
1	1.2375	0.1972	0.7352	0.4834	0.6670	0.5390	0.9477	0.5216	0.5100
2		1.2015	0.7728	0.8330	0.8908	0.9426	0.4662	1.0604	0.6312
3			0.6240	0.4362	0.6107	0.5668	0.9639	0.5337	0.4962
4				0.2670	0.2474	0.4812	0.6496	0.5108	0.2718
5					0.4292	0.2906	0.6542	0.5039	0.1395
6						0.4385	0.6253	0.5479	0.3563
7							0.5162	0.5790	0.3327
8								0.6938	0.5118
9									0.4516

member as well as their alteration type according to Thompson and Thompson (1996). Ten alterations were detected by unmixing of the Hyperion data (Oskouei and Busch 2007). Two alteration maps were then produced with the use of the abundance of each class. In the first map, a pixel is assigned to a class if the respective end-member has maximum abundance in that pixel, and in the second map, the pixel is assigned to an end-member if the maximum abundance is higher than 30%. In this study, these two maps are designated as the maximum map and over-30-% map, respectively. The distribution of each class is selected as the region of interest (ROI) for the subsequent process.

ROIs obtained with the maximum map contain more pixels with a lower probability, of course, and the reverse is true for ROIs obtained with over 30% map. Because of differences between the ASTER and Hyperion ground resolution, the ROIs of classes were also resampled by Reconcile ROI algorithm on ENVI programme to match the pixel size of the ASTER dataset during the registration of the Hyperion dataset on the ASTER scene. To evaluate the ability of ASTER to spectrally discriminate between different classes, their separability was calculated. This option on ENVI computes the spectral separability between selected ROI pairs for a given input file with the use of Jeffries-Mautusita and Transformed Divergence methods (Richards and Jia 2006). The resultant values range from 0.0 to 2.0 and indicate how well the selected ROI pairs are spectrally separated. Values greater than 1.9

indicate that the ROI pairs have good separability. The ROI pairs with lower separability values (less than 1) should be combined into a single ROI (Richards and Jia 2006). First, the separability of ROIs from the maximum map is computed. Table 2 includes the resultant scores for respective pairs of classes. Only three pairs show scores more than 1, and this is not satisfactory. So far, the classification using these classes does not yield accurate results. Fortunately, however, the use of over 30% ROIs as training data resulted in acceptable separability scores, as demonstrated in Table 3. In this case, only classes 1, 3, 6 and 5, and 10 should be combined, as their separability scores are less than 1. This means that we can now classify the ASTER dataset for seven training classes (1–3–6, 2, 4, 5–10, 7, 8, 9). In view of the mineral constituents in the combined classes according to Table 1, 1–3–6 and 5–10 may represent a kind of skarn alteration and calcic skarn, respectively.

Classification

It is obvious that the extension of detected features from Hyperion to ASTER data, whose coverage is much higher than that of Hyperion, will be cost-effective. The identified pixels and their resulted mineralogy from the processing of the Hyperion data of the study area were therefore used as training pixels for classification tasks and then for accuracy

Table 3 Separability computations for over 30% classes

Classes	2	3	4	5	6	7	8	9	10
1	1.9892	0.7857	1.8547	1.4982	0.6965	1.6506	1.8096	1.9864	1.7359
2		1.9836	1.7563	1.7332	1.8658	1.9088	1.5574	1.9991	1.4074
3			1.7604	1.3929	1.7643	1.6458	1.7707	1.9846	1.66863
4				1.1587	1.8225	1.5169	1.5422	1.9889	1.0642
5					1.3380	1.2574	1.3835	1.9748	0.7480
6						1.4163	1.5123	1.9967	1.3343
7							1.0295	1.9967	1.4392
8								1.9975	1.3679
9									1.9893

Table 4 Producer accuracy of classifiers (in percent) for each class using the over 30% map

Classes										
Methods	1–3–6	2	4	5–10	7	8	9	Max	Min	OA
SAM	12.32	71.34	55.27	25.25	62.23	27.88	79.46	79.46	12.32	35.64
ML	32.36	87.01	80.37	29.33	53.78	53.96	95.97	95.97	29.33	47.68
Min-Dist	30.65	80.50	17.70	4.65	35.49	13.48	76.89	80.50	4.65	31.68
Mah-Dist	43.70	75.86	69.50	48.02	66.59	65.02	80.93	80.93	43.70	56.69

analysis by confusion matrix (Oskouei and Busch 2007). The most effective classifier for this purpose could be distinguished with the use of the confusion matrix after classification.

Four classification systems (Spectral Angle Mapper (SAM), maximum likelihood (ML), Mahalanobis distance (Mah-Dist), and minimum distance (Min-Dist)) were then applied to perform the task. The producer accuracy (PA) and overall accuracy (OA) obtained from the confusion matrix were used to evaluate the performance of classification algorithms. The parameters PA and OA are defined as follows:

$$PA_{\omega} = \frac{n_{\omega}}{N_{\omega}} \quad (1)$$

$$OA = \frac{\sum_{i=1}^m n_{\omega i}}{\sum_{i=1}^m N_{\omega i}} \quad (2)$$

Where:

- n_{ω} is the number of pixels classified correctly for class ω by the classifier;
- N_{ω} is the total number of pixels for class ω in the training dataset;
- m is the number of classes.

In Table 4, the PAs of the classifiers are indicated for seven classes; the last column in the table is the OA of each classifier. These are resulted by multiple examinations of the

starting parameters of each classifier to achieve highest possible overall accuracies. In accordance with the result obtained with the confusion matrix, the Mah-Dist classifier yields the best performance for the classification of the ASTER dataset.

On the other hand, the performance of classifiers is not the same for all classes. For instance, the classification of classes 2, 4, and 9 is more accurate by ML than by other methods; and Mah-Dist shows a better functionality for classes 1–3–6, 5–10, 7, and 8. The use of combined classifiers can therefore yield better results. Kittler et al. (1998) developed a common theoretical framework for classifier combination and compared different combination schemes such as the product rule, sum rule, min rule, max rule, median rule, and majority voting. They empirically showed that the sum-rule method yields the best performance. By a sensitivity analysis, they found out that the method is most resilient toward estimation errors; this may provide a plausible explanation for its superior performance.

Other researchers (e.g. Duin 2002 and Lepistö et al. 2004) later showed that the success of combined classifiers depends on different factors, and that there are very few practical instances of their successful performance. In the case of small training sets or overtraining of some base classifiers, the combined classification will result in an unreliable outcome which is dominated by special base classifiers. Two different factors which prevent more accurate results from combined classifications are discussed in the following:

- One reason for this failure is a difference between the statistical distributions of posterior probabilities calculated

Table 5 Mean of posterior probabilities computed by different classifiers for seven classes

Classifiers					
Classes	Min-Dist	SAM	ML	Mah-Dist	
1–3–6	0.442	0.016	0.994	0.020	
2	0.511	0.018	0.992	0.022	
4	0.429	0.017	0.994	0.020	
5–10	0.440	0.017	0.996	0.020	
7	0.436	0.016	0.996	0.020	
8	0.427	0.019	0.996	0.020	
9	0.429	0.016	0.996	0.021	

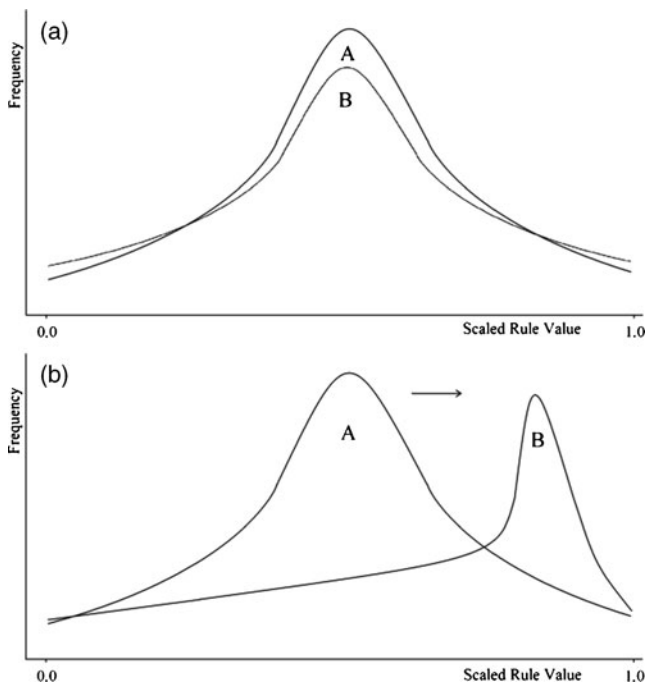


Fig. 2 Posterior probability distribution of two classifiers with **a** similar and **b** different distribution patterns

by base classifiers. Outputs from classifiers are not optimally scaled with respect to one another, because of differences among their models, algorithms, and primary parameters. Scaling of posterior probabilities in the same range will not even be of much help. Note the mean value of posterior probabilities of the four classifiers for the seven classes in Table 5. The means are calculated after scaling of all rule images in the range from 0.0 to 1.0. The mean values of the posterior probabilities for the ML method are very high and close to the maximum (1.0); therefore, this classifier will dominate the outcome of combined classification.

- (b) Another reason is the neglect of differences between computed accuracies for the classes by different classifiers. For example, the most accurate classification for classes 2, 4, and 9 is accomplished by ML; however, for classes 1–3–6, 5–10, 7, and 8, this is done by the Mah-Dist method.

The SCC method presented here is designed to rectify the two above-mentioned detrimental factors.

SCC method

This is a stepwise method that starts by forming a combined rule image (CRI) whose bands are selected from the rule images of the base classifiers with respect to their accuracy. As indicated in Table 5, the rule channels for classes 2, 4, and 9 are the respective channels in the ML rule image; in a

Table 6 Overall accuracies (percent) of three combined classifiers

Classifiers	OA	
	Scaled 0.0–1.0	Equal mean
Sum rule	0.40	2.66
Maximum rule	13.11	0.56
Median rule	0.40	2.66

similar manner, the rule channels for classes 1–3–6, 5–10, 7, and 8 are the respective channels from the Mah-Dist rule image. The use of only the most accurate rule channels will therefore enhance the effectiveness of combined classification from the standpoint of factor b.

Since this method uses only the rules with the highest accuracy, it is necessary to mask unclassified pixels in each rule channel. In this way, the decision will be reached only on common pixels in the final classification. If A and B are presumably statistical distribution of two classes classified by two different classifiers, in the final classification then, the regions A–B and B–A will be assigned to classes A and B, respectively. Therefore, the complementary tasks of the SCC classifier will be decision-making on pixels in the region of $A \cap B$ and their assignment to A or B with regard to their posterior probabilities from two classifiers. As discussed earlier, finding a unique criterion for doing this has always been a challenging issue in combined classification.

The SCC algorithm therefore applies an iterative method to achieve the best possible assignment, which will be examined on the basis of the overall accuracy. The algorithm first scales CRI channels in the range from 0.0 to 1.0; it then classifies scaled CRI and computes OA. This is done based on the following criterion:

$$x \in \omega_j \text{ if } P(\omega_j|x) = \max_{k=1}^n P(\omega_k|x) \tag{3}$$

Where x is the scaled rule value, ω is the class, and n is the number of classes.

If the statistical distributions of posterior probabilities of selected classifiers are similar with a closed mean (e.g.

Table 7 Accuracy computations for SCC after step 4 and final accuracies

Classes	PA, %							OA, %
	1–3–6	2	4	5–10	7	8	9	
After step 4	31.13	86.31	80.37	26.19	58.71	51.15	95.97	47.32
Final results	46.91	77.09	63.79	48.02	66.59	65.02	83.86	57.81

Table 8 Overall accuracy which resulted with the use of the classifiers

Classifiers								
	SAM	ML	Min-Dist	Mah-Dist	Sum rule	Maximum rule	Median rule	SCC
OA, %	35.64	47.68	31.68	56.69	2.66	13.11	2.66	57.81

Fig. 2a), the calculated OA in this step will be an optimum. However, in the case of different distributions like that in Fig. 2b, artefact shifting of one toward another will increase the OA. In this case, the algorithm accomplishes this gradually until the possible maximum OA has been achieved.

The SCC algorithm could be expressed in a stepwise manner as follows:

1. Make CRI
2. Mask unclassified pixels
3. Scale CRI channels (0.0–1.0)
4. Classify CRI

5. Compute overall accuracy (OA1)
6. Add rule channels of low mean classifier(s) by (max-mean-channel mean)/4.0
Maxmean is the maximum of the means of CRI channels
7. Compute overall accuracy (OA2)
8. If $OA2 \leq OA1$ END
9. Go to 6

Results

To achieve the most reliable map of predefined classes in the ASTER scene and also to compare SCC with other combined classifiers, the classifier sum rule, maximum rule, and median rule are tested, in addition to four base classifiers. The functionality of base classifiers has already been discussed in previous sections; their producer and overall accuracies are indicated in Table 5.

Three combined classifiers (sum, maximum, and median (Kittler et al. 1998)) are run with the use of scaled rule images in the range from 0.0 to 1.0. The resulting OAs from these classifiers establishes a very weak functionality, as shown in

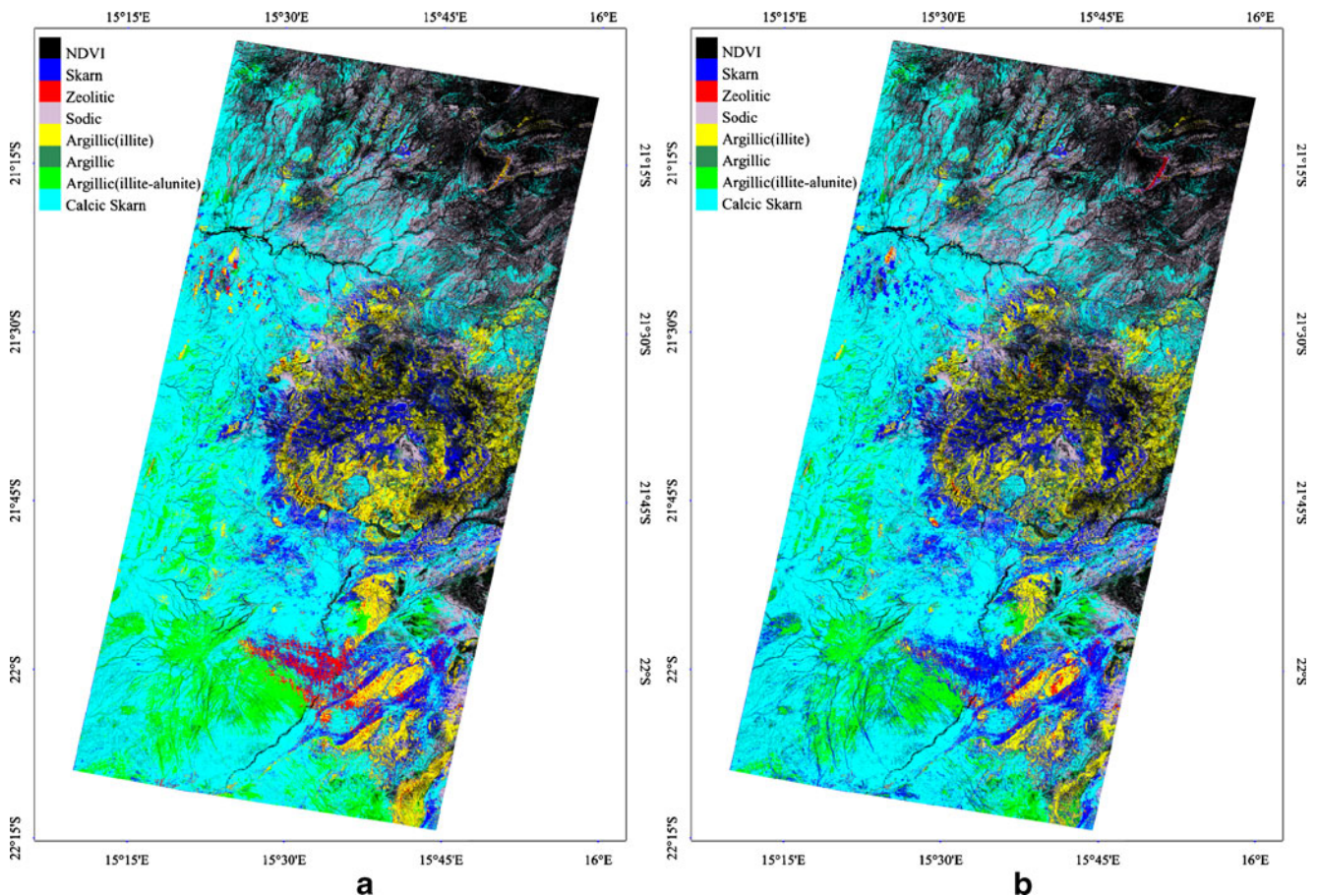
**Fig. 3** The resultant classification maps of detected alterations by Mah-Dist (a) and SCC method (b)

Table 6. Two different methods for the modification of rule images (scaling in the range 0.0–1.0 and equalization of means) are examined, and the respective accuracies are illustrated in the table. As mentioned before, their performance is highly variable from case to case and is strongly dependent on base classifiers and the existence of weak classifiers. Overtraining classification of some base classifiers will also render the combined classification unreliable.

In contrast to the above-mentioned combined classifiers, the outcome of the SCC method is prominent. The results of SCC are illustrated in Table 7. The second row of this table comprises the PAs calculated for seven classes and the OA in the step after scaling of the CRI in the range from 0.0 to 1.0. The third row is the result of the final accuracy computations for SCC.

A comparison of the accuracies which resulted from SCC and three other combined classifiers implies that the functionality of the algorithm presented here is effective and reliable. Furthermore, among all classifiers tested in this study (both base and combined), SCC has yielded the most accurate classification (Table 8). Finally, the alteration thematic maps with the use of Mah-Dist and SCC are illustrated in Fig. 3.

Conclusion

The result from Hyperion dataset unmixing processes is extended to the ASTER scene with the aim of achieving an alterations map in a much broader area. Therefore, the mapped alterations in the Hyperion dataset are used as training datasets for classification of the ASTER scene. At first, the classification of the ASTER data is performed with the use of four classifiers (SAM, ML, Min-Dist, and Mah-Dist). The accuracy evaluation showed that among these the Mah-Dist method yields the highest overall accuracy.

Meanwhile, the capability of combined classification is also investigated, and a new method (SCC) is presented with the aim of achieving the highest possible OA. Other combined classification methods always suffer from two problems. These emanate from the lack of certain criteria for selecting base classifiers, and for modification of posterior probabilities resulted from different base classifiers. The reliability of combined classification functionality, therefore, differs from case to case. The SCC algorithm, however, provides distinct solutions for the two problems and thus always yields a more accurate classification than do base classifiers. The approach presented in this study provided effective performance in comparison with both base and combined classifiers and resulted in the best OA, as

indicated in Table 8. The overall accuracy which results from SCC is always higher than the maximum overall accuracy of base classifiers.

References

- Duin R (2002) The combining classifier: to train or not to train? Proc. of the 16th International Conference on Pattern Recognition (ICPR), Quebec City, Canada, 2002, pp. 765–770
- Favretto A, Geletti R (2004) Satellite imagery elaboration (ASTER sensor, TERRA Satellite), in order to map rock distribution in extreme areas, The Prince Albert Mountain, Victoria Land, Antarctica. Proc. of 20th ISPRS Congress, Istanbul, Turkey, July 12–13, 2004, pp. 1234–1240
- Galvao LS, Almeida-Filho R, Vitorello I (2005) Spectral discrimination of hydrothermally altered materials using ASTER short-wave infrared bands: evaluation in a tropical savannah environment. *Int J Appl Earth Obs Geoinformation* 7:107–114
- Giacinto G, Roli F (2001) An approach to the automatic design of multiple classifier systems. *Pattern Recognit Lett* 22(1):25–33
- Hewson RD, Cudahy TJ, Mizuhiko S, Ueda K, Mauger AJ (2005) Seamless geological map generation using ASTER in the Broken Hill-Curnamona Province of Australia. *Remote Sens Environ* 99:159–172
- Hubbard BE, Crowley JK (2005) Mineral mapping on the Chilean–Bolivian Altiplano using co-orbital ALI, ASTER and Hyperion imagery: data dimensionality issues and solutions. *Remote Sens Environ* 99:173–186
- Kittler J, Hatef M, Duin RPW, Matas J (1998) On combining classifiers. *IEEE Trans Pattern Anal Mach Intell* 20(3):226–239
- Kruse FA (2002) Combined SWIR and LWIR mineral mapping using MASTER/ASTER. Proc. IGARSS 2002, Toronto, Canada, pp. 2267–2269
- Kruse FA, Perry SL, Caballero A (2002) Integrated multispectral and hyperspectral mineral mapping, Los Menucos, Rio Negro, Argentina, Part II: EO-1 Hyperion/AVIRIS Comparisons and Landsat TM/ASTER Extensions. Proc. of 11th JPL Airborne Geoscience Workshop, Jet Propulsion Laboratory, CA, USA, March 4–8, 2002
- Lepistö L, Kunttu I, Autio J, Visa A (2004) Combining classifiers in rock image classification—supervised and unsupervised approach. Proc. of Advanced Concepts for Intelligent Vision Systems, Brussels, Belgium, August 31–September 3, 2004
- Oskouei MM, Busch W (2007) Unmixing and mapping of alteration zones using EO-1 Hyperion data in Erongo, Namibia. Proc. of SPIE Remote Sensing Conference, 17–20 September, Florence, Italy, 2007
- Paskaleva BS, Hayat MM, Moya MM, Fogler RJ (2004) Multispectral rock type separation and classification. Proc. of 49th Annual Meeting of the SPIE: Infrared Spaceborne Remote Sensing XII, Denver, CO, SPIE Proc. 5543, August 2–6, 2004, pp. 152–163
- Richards JA, Jia X (2006) Remote sensing digital image analysis. Springer, Heidelberg
- Rowan LC, Mars JC (2003) Lithologic mapping in the Mountain Pass, California area using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data. *Remote Sens Environ* 84:350–366
- Schneider (2004) Gondwanaland Geopark. The Gondwanaland Geopark Project was made possible with financial support from Unesco (Windhoek Cluster Office), pp. 31–33
- Sharkey AJC, Sharkey NE, Gerecke U, Chandroth GO (2000) The “test and select” approach to ensemble combination. Proc. of the

- First International Workshop on Multiple Classifier Systems (MCS2000), Springer-Verlag, LNCS 1857, pp. 30–44
- Thompson AJB, Thompson JFH (1996) Atlas of alteration. Geological Association of Canada
- Vaughan RG, Hook SJ, Calvin WM, Taranik JV (2005) Surface mineral mapping at steamboat springs, Nevada, USA, with multi wavelength thermal infrared images. *Remote Sens Environ* 99 (1–2):140–158
- Wang X, Zhang S (2006) Evaluation of land cover classification effectiveness for the Queer Mountains, China using ASTER satellite data. *Proc. of Advanced Technology in the Environmental field, Lanzarote, Canary Islands, Spain, 6–8 June, 2006*