

Automatic Martian Dust Storm Detection from Multiple Wavelength Data Based on Decision Level Fusion

KEISUKE MAEDA^{1,a)} TAKAHIRO OGAWA^{1,b)} MIKI HASEYAMA^{1,c)}

Received: March 13, 2015, Accepted: April 20, 2015, Released: July 27, 2015

Abstract: This paper presents automatic Martian dust storm detection from multiple wavelength data based on decision level fusion. In our proposed method, visual features are first extracted from multiple wavelength data, and optimal features are selected for Martian dust storm detection based on the minimal-Redundancy-Maximal-Relevance algorithm. Second, the selected visual features are used to train the Support Vector Machine classifiers that are constructed on each data. Furthermore, as a main contribution of this paper, the proposed method integrates the multiple detection results obtained from heterogeneous data based on decision level fusion, while considering each classifier’s detection performance to obtain accurate final detection results. Consequently, the proposed method realizes successful Martian dust storm detection.

Keywords: Mars, detection, dust storm, decision level fusion

1. Introduction

In recent decades, the human race has pursued the potential for Martian life and has researched Martian environment to emigrate to Mars in the future [1]. Especially, the phenomenon where dust are lifted into the atmosphere is named “dust storm.” Dust storms have a great influence on the Martian environment [2]. Specifically, dust storms have a significant impact on the global temperatures on Mars [3], [4]. Hence, studies on dust storms have attracted much attention. Examples of dust storms are shown in Fig. 1.

In recent studies on Mars, many researchers have received a large amount of Martian data taken by the Mars Orbiter Camera (MOC) on board the Mars Global Surveyor (MGS) [5], [6]. Although dust storm detection from a limited number of Martian data has been previously carried out manually by the researchers [7], [8], many researchers spend a lot of time for detecting dust storms, and it is necessary to detect dust storms automatically and accurately from Martian data. Therefore, it is expected that machine learning could be used for this task. It should be noted that since we utilize Martian data that are more unique and heterogeneous than data utilized in general detection problems, it is difficult to obtain accurate detection results through direct use of a conventional algorithm.

In this paper, we propose a new Martian dust storm detection method from multiple wavelength data based on decision level fusion. As a main contribution of this paper, we try to integrate

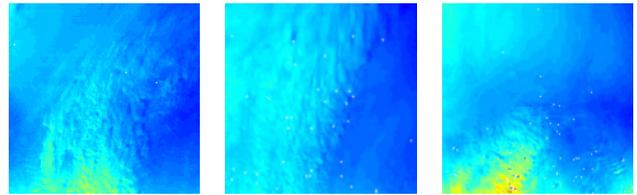


Fig. 1 Examples of dust storms.

detection results from multiple wavelength data by using decision level fusion [9], [10], [11], [12]. First, the proposed method extracts visual features from heterogeneous data and selects optimal features based on the minimal-Redundancy-Maximal-Relevance (mRMR) algorithm [13]. Then the selected visual features are used to train Support Vector Machine (SVM) [14] classifiers that are constructed on each wavelength data. Furthermore, it is necessary to integrate the obtained multiple detection results effectively while considering each classifier’s performance in order to improve dust storm detection. Therefore, we adopt decision level fusion [9], [10], [11], [12], which assigns higher weights to the detection results of the best classifiers and enables the successful integration of multiple detection results.

2. Dust Storm Detection

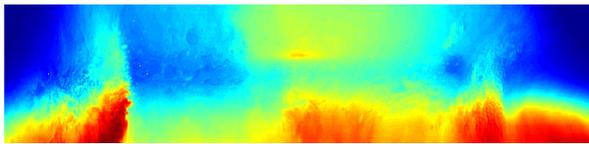
Automatic detection of dust storms that utilizes multiple wavelength data are presented in this section. In Section 2.1, we provide an overview of Martian data. In Section 2.2, we extract visual features from the multiple wavelength data and select the optimal visual features. In Section 2.3, we generate an SVM classifier on each data. In Section 2.4, we integrate multiple detection results obtained from the SVM classifiers based on the decision level fusion.

¹ Hokkaido University, Sapporo, Hokkaido 060–0814, Japan

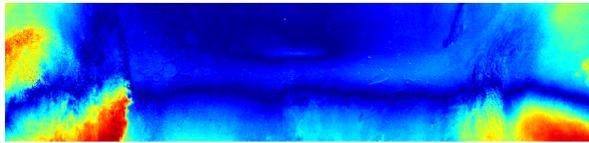
a) maeda@lmd.ist.hokudai.ac.jp

b) ogawa@lmd.ist.hokudai.ac.jp

c) miki@ist.hokudai.ac.jp



(a) RED reflectance data



(b) Background subtraction data obtained from RED reflectance data

Fig. 2 Martian data used in our detection method. Each data has one value representing the reflectance value at each axis, i.e., each pixel. The right side corresponds to the north polar cap. In (a) and (b), the values become larger from blue to red. Each data has a resolution of $384 \times 1,600$ pixels.

2.1 Martian Data of MOC

In this subsection, an overview of the Martian data are presented. In studies on Mars, many researchers have utilized the reflectance data since only the reflectance data have been sent from the satellites. Since the MOC had obtained daily global maps of the Martian surface in two wavelength bands: BLUE (400 nm–450 nm) and RED (575 nm–625 nm), we utilize both BLUE and RED reflectance data (Fig. 2 (a)).

Since dust storms occur sporadically on Mars, it is effective to utilize subtraction data (Fig. 2 (b)) from background data (Fig. 3) that does not include dust storms. Note that Fig. 3 shows the whole surface of Mars, and Fig. 2 (a) shows only a part of the surface of Mars. Therefore, we automatically generate the background subtraction data and introduce them into our detection method. Accordingly, the proposed method utilizes four kinds of data that are taken in the same area, i.e., RED and BLUE reflectance data and background subtraction data of RED and BLUE reflectance data.

In order to generate background subtraction data, we first generate background data that do not include any dust storms from the reflectance data. From information about the latitude and longitude obtained with the reflectance data, we performed voting on the reflectance values according to the corresponding axis on the Martian map. Then, by calculating their median values for each axis, background data can be obtained. Since dust storms occur sporadically, their reflectance values tend not to be selected by calculating the median values. Therefore, we can obtain background data from the reflectance value by not including any dust storms as shown in Fig. 3. Then, we take the difference between the reflectance data and the background data to generate the background subtraction data.

2.2 Extraction and Selection of Visual Features from Multiple Wavelength Data

In this subsection, the extraction and selection of visual features are presented. Since the proposed method focuses on the difference in local characteristics to detect dust storms, we utilize the HOG feature [16] and the DSIFT feature [17] as local visual

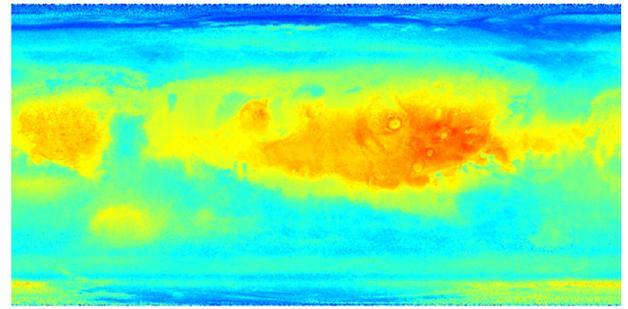


Fig. 3 Background data on Mars calculated from RED reflectance data. The upper and lower sides respectively correspond to the north polar cap and the south polar cap. In this figure, the values become larger from blue to red. Background data has a resolution of $1,800 \times 3,600$ pixels.

features. Thus, we can obtain a feature vector of $D (= 9,180)$ ^{*1} dimensions from each data. This feature vector is calculated from the k -th ($= 1, \dots, k_{\max}$) kind of data and defined as $\mathbf{v}_k \in \mathbb{R}^D$, such that each element of \mathbf{v}_k is represented as $v_k(j)$ ($j = 1, 2, \dots, D$). Note that as shown in the previous section, we use the four data and $k_{\max} = 4$ in this paper. For each data, the label $y \in \{1, -1\}$ denotes the existence or nonexistence of dust storms. Then $d_k (< D)$ optimal features are selected from D features based on the mRMR algorithm [13] as a new feature vector $\gamma_k \in \mathbb{R}^{d_k}$. Therefore, the proposed method can remove features that caused performance degradation to obtain features that have close relevance to dust storm detection.

2.3 Generation of SVM Classifier for Each Data

In this subsection, the generation of the SVM classifier for each different kind of data is presented. Given the k -th training data consisting of vectors $\gamma_k^i \in \mathbb{R}^{d_k}$ ($i = 1, 2, \dots, N$; N being the number of training data) and their corresponding true class labels $y^i \in \{1, -1\}$, an optimal SVM hyperplane is calculated by using the training dataset. Furthermore, the class label of the k -th test data $\gamma_k \in \mathbb{R}^{d_k}$ is obtained as

$$f_k(\gamma_k) = \text{sign}\left(\sum_{i=1}^N c_k^i y^i K_k(\gamma_k^i, \gamma_k) + e_k\right), \quad (1)$$

where $K_k(\cdot, \cdot)$ is the kernel function, c_k^i is the Lagrange multiplier, and e_k is the constant value. By using the SVM classifier, we can perform dust storm detection for each k -th kind of data.

2.4 Integration of Multiple Detection Results

This subsection presents the integration of the multiple detection results obtained from the SVM classifiers based on the decision level fusion. Since the method in Ref. [9] has come from the research field of computer-aided diagnosis (CAD), it integrates multiple classification results from each human annotator, e.g., radiologist. We regard the k_{\max} SVM classifiers as k_{\max} annotators. In order to integrate multiple results, we focus on the detection performance of each annotator and assign higher weights to results of annotators which have higher detection performance. An overview of the target model and its training and test phases is shown below.

^{*1} We can obtain 2,268 -dimensional values for the HOG feature and 6,912 -dimensional values for the DSIFT feature.

2.4.1 Performance of Each Annotator and Classification Model

The performance α_k (sensitivity) and β_k (specificity) of the k -th SVM classifier are defined as follows:

$$\alpha_k = \Pr[y_k = 1 | y = 1], \quad (2)$$

$$\beta_k = \Pr[y_k = 0 | y = 0], \quad (3)$$

where the binary class label $y_k \in \{1, 0\}$ are assigned to the feature vector γ_k by the k -th annotator, and the binary true class label $y \in \{1, 0\}$ is the Ground Truth. Note that y_k and y are assigned by respectively rewriting $y_k \in \{1, -1\}$ and $y \in \{1, -1\}$ in the previous subsection.

The proposed method adopts a linear discriminating function whose classification model is specifically written as

$$f_w(\mathbf{x}) = \mathbf{w}^T \mathbf{x}, \quad (4)$$

where \mathbf{w} is a parameter and \mathbf{x} is a feature vector of the test data. Note that \mathbf{x} is obtained by aligning features selected from $\gamma_1, \dots, \gamma_{k_{\max}}$ based on the mRMR algorithm, where $\gamma_1, \dots, \gamma_{k_{\max}}$ are the feature vectors of k_{\max} kinds of data.

2.4.2 Training Phase

Given the training data $\mathfrak{Y} = \{y^i, \mathbf{x}^i, y_1^i, \dots, y_{k_{\max}}^i\}_{i=1}^N$ containing N instances, where y^i is the true class label of the i -th training data. The probability p^i for the true positive class is modeled as

$$p^i = \Pr[y^i = 1 | \mathbf{x}^i, \mathbf{w}] \quad (5)$$

$$= \psi(\mathbf{w}^T \mathbf{x}^i) \quad (6)$$

$$= \frac{1}{1 + \exp(-\mathbf{w}^T \mathbf{x}^i)}. \quad (7)$$

Given the training data \mathfrak{Y} , the parameter \mathbf{w} , $\boldsymbol{\alpha} = \{\alpha_1, \dots, \alpha_{k_{\max}}\}$, $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_{k_{\max}}\}$, the conditional likelihood is defined as follows:

$$\begin{aligned} \Pr[\mathfrak{Y} | \mathbf{w}] &= \prod_{i=1}^N \Pr[y_1^i, \dots, y_{k_{\max}}^i | y^i = 1, \boldsymbol{\alpha}] \cdot \Pr[y^i = 1 | \mathbf{x}^i, \mathbf{w}] \\ &+ \prod_{i=1}^N \Pr[y_1^i, \dots, y_{k_{\max}}^i | y^i = 0, \boldsymbol{\beta}] \cdot \Pr[y^i = 0 | \mathbf{x}^i, \mathbf{w}]. \end{aligned} \quad (8)$$

By assuming that $y_1^i, \dots, y_{k_{\max}}^i$ are independent, the likelihood of Eq. (8) is rewritten as

$$\Pr[\mathfrak{Y} | \mathbf{w}] = \prod_{i=1}^N [a^i \psi(\mathbf{w}^T \mathbf{x}^i) + b^i (1 - \psi(\mathbf{w}^T \mathbf{x}^i))], \quad (9)$$

$$a^i = \prod_{k=1}^{k_{\max}} [\alpha_k]^{y_k^i} [1 - \alpha_k]^{1-y_k^i}, \quad (10)$$

$$b^i = \prod_{k=1}^{k_{\max}} [1 - \beta_k]^{y_k^i} [\beta_k]^{1-y_k^i}. \quad (11)$$

We can find the optimal parameter \mathbf{w} by maximizing the log-likelihood as follows:

$$\hat{\mathbf{w}} = \arg \max_{\mathbf{w}} \{\ln \Pr[\mathfrak{Y} | \mathbf{w}]\}. \quad (12)$$

In Ref. [9], the log-likelihood is maximized by the Expectation-Maximization (EM) algorithm [18]. Let $\mathbf{y} = [y^1, \dots, y^N]$ be the set of the actual labels, and the complete data log-likelihood is written as follows:

$$\ln \Pr[\mathfrak{Y}, \mathbf{y} | \mathbf{w}] = \sum_{i=1}^N [y^i \ln(p^i a^i) + (1 - y^i) \ln\{(1 - p^i) b^i\}]. \quad (13)$$

In order to obtain the optimal parameter \mathbf{w} , the EM algorithm alternately calculates the following two steps.

(1) E-step:

Given the training data \mathfrak{Y} and the current estimate of the model parameter \mathbf{w} , the conditional expectation is calculated as

$$\mathbb{E}[\ln \Pr[\mathfrak{Y}, \mathbf{y} | \mathbf{w}]] = \sum_{i=1}^N [\mu^i \ln(p^i a^i) + (1 - \mu^i) \ln\{(1 - p^i) b^i\}], \quad (14)$$

where $\mu^i = \Pr[y^i = 1 | y_1^i, \dots, y_{k_{\max}}^i, \mathbf{x}^i, \mathbf{w}]$. By using the Bayesian theorem, the probabilistic label μ^i is obtained as follows:

$$\mu^i \propto \Pr[y_1^i, \dots, y_{k_{\max}}^i | y^i = 1, \mathbf{w}] \cdot \Pr[y^i = 1 | \mathbf{x}^i, \mathbf{w}] \quad (15)$$

$$= \frac{a^i \psi(\mathbf{w}^T \mathbf{x}^i)}{a^i \psi(\mathbf{w}^T \mathbf{x}^i) + b^i (1 - \psi(\mathbf{w}^T \mathbf{x}^i))}. \quad (16)$$

(2) M-step:

In this step, we estimate the parameter \mathbf{w} from the current estimated μ^i and the training data \mathfrak{Y} . Due to the non-linearity of the sigmoid function, there is no closed form solution for \mathbf{w} . Therefore, the parameter \mathbf{w} is estimated based on the Newton-Raphson method [19] as follows:

$$\mathbf{w} \leftarrow \mathbf{w} - \eta \mathbf{H}^{-1}(\mathbf{w}) \mathbf{g}(\mathbf{w}), \quad (17)$$

$$\mathbf{g}(\mathbf{w}) = \sum_{i=1}^N \{\mu^i - \psi(\mathbf{w}^T \mathbf{x}^i)\} \mathbf{x}^i, \quad (18)$$

$$\mathbf{H}(\mathbf{w}) = - \sum_{i=1}^N \psi(\mathbf{w}^T \mathbf{x}^i) \{1 - \psi(\mathbf{w}^T \mathbf{x}^i)\} \mathbf{x}^i (\mathbf{x}^i)^T. \quad (19)$$

In the above equations, $\mathbf{g}(\mathbf{w})$ is the gradient vector, $\mathbf{H}(\mathbf{w})$ is the Hessian matrix, and η is a step length. By iterating the above steps until the result of \mathbf{w} converges, the optimal parameter \mathbf{w} and μ^i are obtained.

2.4.3 Test Phase

Given the test data, the final classification result can be obtained as follows. In the previous phase, we essentially solved a regular logistic regression problem with probabilistic labels μ^i . Thus, the final classification result can be obtained by applying a threshold to μ calculated from the test data $\{\mathbf{x}, y_1, \dots, y_{\max}\}$. The value of μ is computed by using p , a and b calculated from the training data. Specifically, $p = \frac{1}{1 + \exp(-\mathbf{w}^T \mathbf{x})}$ is calculated. Furthermore, $a = \prod_{k=1}^{k_{\max}} [\alpha_k]^{y_k} [1 - \alpha_k]^{1-y_k}$ and $b = \prod_{k=1}^{k_{\max}} [1 - \beta_k]^{y_k} [\beta_k]^{1-y_k}$ are obtained, where α_k and β_k are respectively sensitivity and specificity of annotator k calculated from the training data and y_k is a classification result of the test data. Since we can calculate $\mu = \frac{ap}{ap + b(1-p)}$ by using p , a and b , the final classification result is obtained as follows:

$$y = \begin{cases} 1 & \text{if } \mu > \lambda \\ 0 & \text{otherwise.} \end{cases}$$

The dust storms of the test data can be detected when $\mu > \lambda$, where λ is a predefined threshold.

Table 1 Details of each dust storm and number of correctly detected dust storms. Since a dust storm is sequentially taken by MOC, there are several data of the dust storms for each Ls.

Ls	Central		Area ($\times 10^6$ km 2)	N_a
	Latitude	Longitude		
161.47	-30.88	149.88	1.9	13/14
163.07	74.63	322.88	5.6	44/45
202.53	51.70	16.51	6.1	4/16
209.47	15.88	45.00	2.4	11/19
221.225	21.96	34.05	3.9	14/16
226.227	-59.50	134.25	3.5	11/11

Table 2 Recall, Precision and F-measure obtained in the experiment.

	Recall	Precision	F-measure
Comparative method 1	0.74	0.77	0.75
Comparative method 2	0.73	0.74	0.74
Comparative method 3	0.70	0.80	0.75
Comparative method 4	0.80	0.77	0.78
Comparative method 5	0.78	0.81	0.79
Proposed method	0.80	0.82	0.81

3. Experiment Results

This section shows experimental results to verify the effectiveness of the proposed method. Total of 312 data including 121 existence data and 191 nonexistence data are utilized in the experiment. In our experiments, we performed detection of “regional dust storms”^{*2}. The verification method used was 6-fold cross-validation. **Table 1** shows the data divided into 6 groups based on Ls in the verification, where Ls is the areocentric longitude of the Sun along the Martian ecliptic. Central Latitude and Longitude are the central location of the dust storms on Mars. Area represents the region of the dust storms. The dimension d_k of the feature vector selected by the mRMR algorithm is set to the values which output the best SVM classification results. Furthermore, the dimension of x^i based on the mRMR algorithm is set to the values which output the final classification results. In this paper, we used the Gaussian kernel for the kernel function in the SVM with the kernel parameter obtained via grid search [20]. Furthermore, each data has a resolution of $384 \times 1,600$ pixels.

In order to verify the effectiveness of the proposed method, we compare its detection performance with those of the following five methods.

- Comparative method 1 is a method that obtains the results by only using the RED reflectance data.
- Comparative method 2 is a method that obtains the results by only using the BLUE reflectance data.
- Comparative method 3 is a method that obtains the results by only using the background subtraction data (RED).
- Comparative method 4 is a method that obtains the results by only using the background subtraction data (BLUE).
- Comparative method 5 is a method that integrates the multiple detection results by using majority voting^{*3}.

For evaluation, we utilized Recall, Precision and F-measure. The results are shown in **Table 2**. In this table, recall, precision and F-measure, which is the harmonic mean of Recall and Preci-

^{*2} Dust storms are categorized into “local dust storms” ($< 1.6 \times 10^6$ km 2) and “regional dust storms” ($> 1.6 \times 10^6$ km 2) that lasted more than two Martian days) in Ref. [4].

^{*3} In the experiment, majority voting obtains a positive pattern when more than two classifiers agree on a positive pattern.

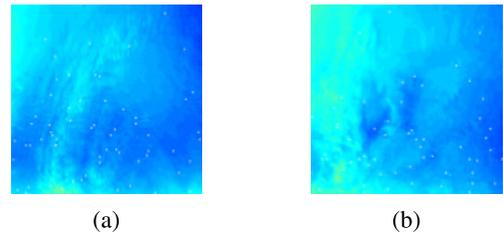


Fig. 4 The data that the proposed method detects incorrectly: (a) overdetection and (b) misdetection.

sion, is the highest for the proposed method. Therefore, the proposed method based on the decision level fusion can detect dust storms automatically and accurately to outperform the comparative methods. In addition, Table 1 also shows the experimental results of the proposed method. Note that N_a is the detection ratio of the dust storms. In this table, the proposed method can detect dust storms in various sizes (Areas). On the other hand, the result of the third row (Ls = 202.53) is worse since classifiers are trained from small dust storms due to the cross validation scheme. Thus, the classifier incorrectly tends to detecting the largest dust storms. Nevertheless, perfect detection of all dust storms is quite a difficult task. This is because some dust storms have only small differences with Martian surfaces. **Figure 4** shows the data that the proposed method detects incorrectly. Figure 4(a) shows an example of overdetection, and Fig.4(b) shows an example of misdetection. Dust storm detection is quite a difficult task since it has only small characteristics. Nevertheless, since the proposed method can detect them more accurately than the comparative methods, its effectiveness can be verified.

4. Conclusions

This paper has presented automatic Martian dust storm detection from multiple wavelength data. In order to improve detection performance, we introduce the decision level fusion integrating multiple results obtained from heterogeneous data into the proposed method. The experimental results show its effectiveness and verify that the proposed method realizes successful detection of Martian dust storms.

When the performances of multiple classifiers are different from each other, the decision level fusion shown in this paper enables successful integration in not only Martian data but also other different data. It will be reported in subsequent work.

Acknowledgments In this research, we utilized Martian data that was kindly provided by JAXA. We would also like to sincerely thank Kazunori Ogohara and Japan Aerospace eXploration Agency (JAXA), for providing Martian data as well as invaluable advice. This work was partly supported by Grant-in-Aid for Scientific Research (B) 25280036 from JSPS.

References

- [1] Shang, C. and Barnes, D.: Fuzzy-rough feature selection aided support vector machines for mars image classification, *Comput. Vision Image Understand*, Vol.117, No.3, pp.202–213 (2013).
- [2] Wang, H., Richardson, M.I., Wilson, R.J., Ingersoll, A.P., Toigo, A.D. and Zurek, R.W.: Cyclones, tides, and the origin of a cross-equatorial dust storm on Mars, *Geophys. Res. Lett.*, Vol.30, no.9, 1488, DOI: 10.1029/2002GL016828 (2003).
- [3] Strausberg, M.J., Wang, H., Richardson, M.I., Ewald, S.P. and Toigo, A.D.: Observations of the initiation and evolution of the 2001

- Mars global dust storm, *Journal of Geophysical Research*, Vol.110, No.E02006 (2005).
- [4] Cantor, B.A., James, P.B., Caplinger, M. and Wolff, M.J.: Martian dust storms: 1999 Mars Orbiter Camera observations, *Journal of Geophysical Research*, Vol.106, No.E10, pp.23653–23687 (2001).
- [5] Edgett, K.S. and Malin, M.C.: New views of Mars eolian activity, materials, and surface properties: Three vignettes from the Mars Global Surveyor Mars Orbiter Camera, *Journal of Geophysical Research: Planets (1991–2012)*, Vol.105, No.E1, pp.1623–1650 (2000).
- [6] Malin, M.C. and Edgett, K.S.: Mars global surveyor Mars orbiter camera: interplanetary cruise through primary mission, *Journal of Geophysical Research: Planets (1991–2012)*, Vol.106, No.E10, pp.23429–23570 (2001).
- [7] Li, R., Di, K. and Xu, F.: Automatic Mars landing site mapping using surface-based images, *ISPRS WG IV/9: Extraterrestrial Mapping Workshop on Advances in Planetary Mapping*, Vol.22 (2003).
- [8] Castano, A., Fukunaga, A., Biesiadecki, J., Neakrase, L., Whelley, P., Greeley, R., Lemmon, M., Castano, R. and Chien, S.: Automatic detection of dust devils and clouds on Mars, *Machine Vision and Applications*, Vol.19, No.5-6, pp.467–482 (2008).
- [9] Raykar, V.C., Yu, S., Zhao, L.H., Jerebko, A., Florin, C., Valadez, G.H., Bogoni, L. and Moy, L.: Supervised learning from multiple experts: whom to trust when everyone lies a bit, *Proc. 26th Annual International Conference on Machine Learning*, pp.889–896 (2009).
- [10] Kawakami, T., Ogawa, T. and Haseyama, M.: Vocal segment estimation in music pieces based on collaborative use of EEG and audio features, *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp.1197–1201 (2013).
- [11] Kawakami, T., Ogawa, T. and Haseyama, M.: Novel image classification based on decision-level fusion of EEG and visual features, *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp.5915–5919 (2014).
- [12] Ishihara, K., Ogawa, T. and Haseyama, M.: Helicobacter pylori infection detection from multiple X-ray images based on decision level fusion, *IEEE International Conf. Image Processing (ICIP)*, pp.2769–2773 (2014).
- [13] Peng, H., Long, F. and Ding, C.: Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 27, No.8, pp.1226–1238 (2005).
- [14] Cortes, C. and Vapnik, V.: Support-Vector Networks, *Machine Learning*, Vol.20, pp.273–297 (1995).
- [15] Wang, H. and Ingersoll, A.P.: Martian clouds observed by Mars Global Surveyor Mars Orbiter Camera, *J. Geophys. Res.*, Vol.107, No.E10, pp.8-1–8-6, DOI: 10.1029/2001JE001815 (2002).
- [16] Bandeira, L., Marques, J.S., Saraiva, J. and Pina, P.: Automated detection of Martian dune fields, *Geoscience and Remote Sensing Letters, IEEE*, Vol.8, No.4, pp.626–630 (2011).
- [17] Lin, Y., Bunte, M., Saripalli, S., Bell, J. and Greeley, R.: Autonomous volcanic plume detection on planetary bodies, *Acta Astronautica*, Vol.97, pp.151–163 (2014).
- [18] Dempster, A.P., Laird, N.M. and Rubin, D.B.: Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B*, Vol.39, No.1, pp.1–38 (1977).
- [19] Israel, A.B.: A Newton-Raphson method for the solution of systems of equations, *Journal of Mathematical Analysis and Applications*, Vol.15, No.2, pp.243–252.
- [20] Hsu, C.W., Chang, C.C. and Lin, C.J.: A practical guide to support vector classification, Technical Report, Department of Computer Science, National Taiwan University (2003).

(Communicated by Takeshi Masuda)