PAPER *Special Section on Smart Multimedia & Communication Systems*

Classifying Insects from SEM Images Based on Optimal Classifier Selection and D-S Evidence Theory

Takahiro OGAWA[†a)], ***Member***, **Akihiro TAKAHASHI**[†], ***Nonmember***, ***and*** **Miki HASEYAMA**[†], ***Member***

SUMMARY In this paper, an insect classification method using scanning electron microphotographs is presented. Images taken by a scanning electron microscope (SEM) have a unique problem for classification in that visual features differ from each other by magnifications. Therefore, direct use of conventional methods results in inaccurate classification results. In order to successfully classify these images, the proposed method generates an optimal training dataset for constructing a classifier for each magnification. Then our method classifies images using the classifiers constructed by the optimal training dataset. In addition, several images are generally taken by an SEM with different magnifications from the same insect. Therefore, more accurate classification can be expected by integrating the results from the same insect based on Dempster-Shafer evidence theory. In this way, accurate insect classification can be realized by our method. At the end of this paper, we show experimental results to confirm the effectiveness of the proposed method.

key words: *scanning electron microphotograph, insect classification, grouping scheme, result integration*

1. Introduction

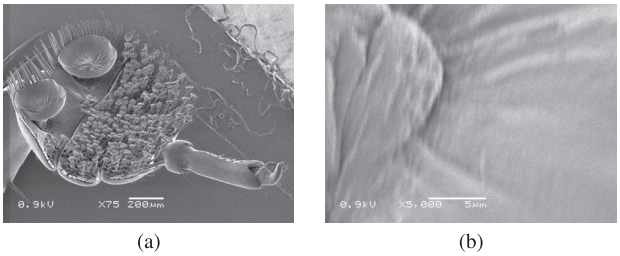
There are a large number of organisms on Earth, and it is estimated that the total number of species in the world varies from 5 million to over 50 million [1]. The diversity of these organisms (plants, animals, insects and other living things) is called biodiversity, and in order to investigate biodiversity, biologists classify organisms by focusing on their shapes in the present taxonomy. However, the classification results may be different among biologists since each biologist uses different criteria for the classification. DNA barcoding*[2] has attracted attention as a new approach for investigating biodiversity, and it provides a solution for the above-described problem. However, it is difficult to apply classification using DNA information to a large number of organisms.

On the other hand, with the development of the scanning electron microscope (SEM), it has become possible to observe organisms' microstructures that are invisible to the naked eye. By utilizing an SEM, biologists have attempted to reveal the differences of microstructures between species. In the field of biomimetics [3], researchers have discovered the relationships between microstructures of organisms and their functions by monitoring images of organisms' surfaces taken by an SEM. Furthermore, according to the discovered relationships, researchers have applied organisms' microstructures to the development of nanomaterials, nanodevices and processes. Therefore, application of SEM images to engineering has become a focus of attention. Due to this situation, the number of SEM images has been increasing recently. Images of insects are more frequently taken by an SEM than other organisms since they have various functions. Thus, we focus on SEM images of insects in this paper.

However, if biologists manually check many images by visual observations, classification requires too much time and labor. Therefore, a method that can classify insects from SEM images automatically and accurately is desirable.

In recent years, rapid progress has been made in the field of visual category recognition. In this field, new features and new classifiers, which are suitable for visual category recognition, have been proposed. Organism classification methods that include calculation of visual features and construction of classifiers have also been proposed [4]–[10]. However, there has been no proposal of a classification method for SEM images. Details of works related to our study are shown in the following section.

Even if target insects for which images are taken by an SEM are the same, visual characteristics, i.e., visual features, are different in magnified images as shown in Figs. 1(a) and (b). Therefore, if conventional methods are directly applied to these visually different images, inaccurate classification results will be obtained. In order to successfully classify these images, it becomes necessary to construct several classifiers for each magnification. Unfortunately, no method that can realize such a classification



(a) (b)

Fig. 1 Examples of SEM images with different magnifications: (a) SEM images of Eretes sticticus at the magnification ×75, (b) SEM images of Eretes sticticus at the magnification ×5000.

*DNA barcoding is a classification method that uses a short genetic marker in an organism's DNA.

Manuscript received February 8, 2016.
Manuscript revised June 11, 2016.

[†]The authors are with the Graduate School of Information Science and Technology, Hokkaido University, Sapporo-shi, 060-0814 Japan.
a) E-mail: ogawa@lmd.ist.hokudai.ac.jp

DOI: 10.1587/transfun.E99.A.1971

scheme has been proposed.

In this paper, we present a new method for insect classification using SEM images. In the proposed method, an optimal training dataset is generated for each magnification. Specifically, we divide a training image dataset including several patterns of magnifications into several groups of images taken with different neighboring magnifications. In this approach, the classification performance obtained by a leave-one-out scheme [11] is monitored using a classifier constructed from each candidate group, and the optimal group, which provides the highest performance, is selected as the final training grouped dataset of each magnification. Then, based on the optimal classifier constructed from the optimal grouped training dataset, accurate classification of insects becomes feasible. In addition, we adopt a new procedure for performance improvement based on a unique characteristic of SEM images. Generally, for each insect, several images are taken by an SEM with different magnifications. Therefore, more accurate classification can be expected by integrating several results obtained from the same sample. In [12]–[14], Dempster-Shafer (D-S) evidence theory [15] is used as a good technology of decision level fusion. Therefore, in the proposed method, we integrate the classification results based on D-S evidence theory. Owing to the above novel approaches, the proposed method realizes accurate insect classification based on unique characteristics of SEM images. This paper is an extended version of [16]. By introducing the new integration scheme, we improve the classification performance in this paper.

This paper is organized as follows. In Sect. 2, we introduce works related to this study. In Sect. 3, we explain the proposed method. Section 4 shows experimental results to verify the effectiveness of the proposed method. Finally, conclusions are given in Sect. 5.

2. Related Work

In this section, works related to our study are presented. Specifically, we show brief reviews of visual feature extraction, general image classification and organism image classification in 2.1, 2.2 and 2.3, respectively.

2.1 Review of Visual Feature Extraction

The most representative local feature in recent years was Scale Invariant Feature Transform (SIFT) proposed by Lowe [17], and its variants such as Principal Component Analysis (PCA)-SIFT [18] and Background and SIFT (BSIFT) [19] were also proposed. Furthermore, in order to reduce their computation complexity with keeping high image representation ability, Speeded Up Robust Features (SURF) [20] were proposed. The Bag-of-Features (BoF) representation [21] enables successful extraction of visual features from general images using the above local features, and it has been one of the most widely used features. In addition, there have been proposed many extensions of the BoF method [22]–[24]. Furthermore, several different approaches have been proposed in recent years. Histograms of Oriented Gradients (HOG), a feature descriptor for a localized region, was proposed [25], and it has been widely used in generic object recognition.

With the rapid growth of deep neural networks [26], Convolutional Neural Network (CNN) features have been developed [27]. It is reported that by using outputs of middle layers of CNN as visual features, the classification performance can be improved in the field of generic object recognition.

In this paper, our main focus is the proposition of the new classification approach of SEM images not introduction of new features. Therefore, we simply use the benchmarking features, SURF-BoF features and HOG features, in the proposed method. Their details are shown in 3.1.

2.2 Review of General Image Classification

This subsection shows the brief review of general image classification. We show some examples of the previously reported classification methods.

Zhang et al. [28] proposed a classification method that combines Support Vector Machine (SVM) [29] and k-Nearest Neighbor (k-NN) [30]. Their method finds neighbors close to a query sample and trains a local SVM that preserves the distance function on the collection of neighbors. Talaat et al. [31] extended the penalized likelihood classification to a multiclass case. Their method utilizes a penalty term based on k-NN and the likelihood of the training patterns' classifications. Kumar and Gopal [32] proposed an improved version of a one-against-all method for multiclass SVM classification based on a decision tree approach. Their method increases the classification speed of a one-against-all method by using posterior probability estimates of binary SVM outputs. McCann and Lowe [33] proposed an improvement of the Naive Bayes Nearest Neighbor (NBNN) image classification algorithm. Their method merges all of the reference data together into one search structure, allowing quick identification of a descriptor's local neighborhood. Bo et al. [34] proposed a novel feature learning architecture. Their method combines a collection of hierarchical sparse features for image classification to capture multiple aspects of discriminative structures.

Recently, there have been proposed a number of novel and attractive image classification methods. Especially, with the rapid growth of deep learning technologies [26], CNN realizes drastic improvement of image classification performance in the field of generic object recognition [35]. On the other hand, it is well known that a large number of training images are generally required for these technologies.

Although CNN can be used as a high performance classifier, target images in this paper are SEM images whose characteristics are quite different from those of generic object recognition. Furthermore, it is difficult in our study to collect a number of training images. Therefore, in the proposed method, we adopt SVM for its generalization characteristic and focus on the introduction of the new classifica-

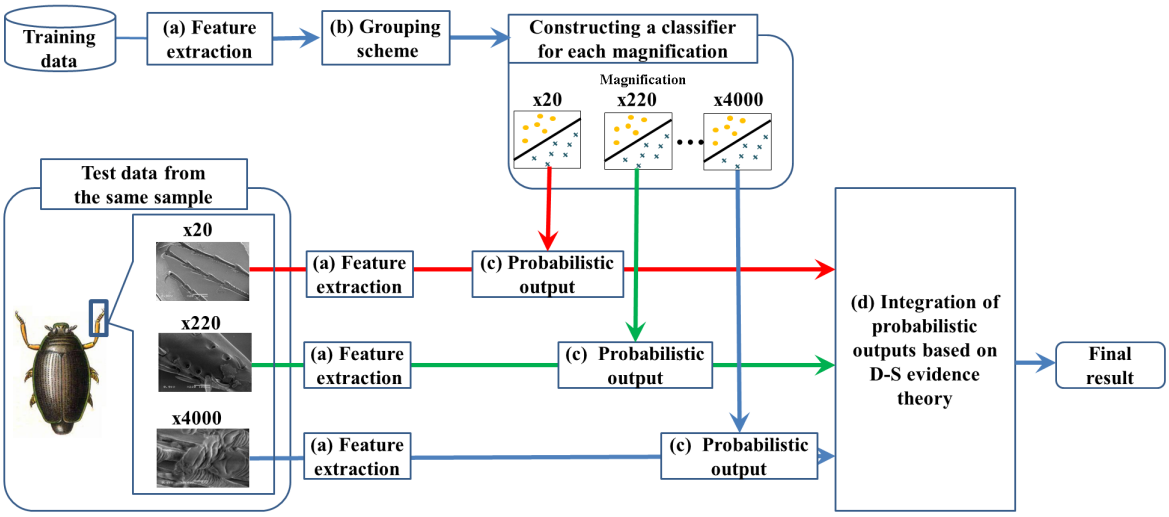


Fig. 2 An overview of the proposed method. (a) Feature extraction is described in 3.1. (b) Grouping scheme is explained in 3.2. (c) Probabilistic output is shown in 3.3. (d) Integration based on D-S evidence theory is described in 3.4.

tion approach that is matched to the characteristics of SEM images.

2.3 Review of Organism Image Classification

Rsearchers have also proposed methods that apply visual recognition methods to images of organisms. Zhang et al. [4] proposed a method that classifies storedproduct insects. Their method classifies insect images based on AdaBoost [36] with a backpropagation neural network [37] to improve the performance. Lu et al. [5] proposed a hybrid approach which combines local and discriminative coding strategies. This method obtains the vector representation of images via spatial pyramid pooling of patches and then classifies images into species by using a linear SVM. Shi-Guo et al. [6] proposed an insect recognition method based on SURF [20]. Local features of insect images are extracted and then the features are utilized as input of a multi-scale histogram algorithm for object matching. Xiao-Lin et al. [7] proposed an insect recognition method based on spectral regression linear discriminant analysis (LDA). Their method utilizes spectral regression LDA to reduce high-dimensional spaces of insect images. Then this method obtains an insect feature subspace and recognizes species by the k-NN algorithm. Le-Qing and Zhen [8] proposed a method to classify insects by analyzing color histograms and Gray-Level Co-occurrence Matrices (GLCM) of wing images. Their method extracts color features and GLCM features from these images, and their matching is performed. In recent years, Venugoban and Ramanan proposed image classification of paddy field insects based on gradient-based features such as SIFT, SURF and HOG features, and they also reported that the combination of SURF+HOG features provided the best classification performance [9]. Furthermore, Shraddha et al. proposed an automatic identification method of agriculture pest insects by using simple color and texture features and an SVM classifier [10].

As shown in the above brief review, since it is difficult to collect a number of training samples in the field of organism image classification, the previous works tended to use benchmarking features and classifiers that have generalization characteristics. Next, there is no study focusing on the insect classification from SEM images. As described above, SEM images have the unique characteristics, and it is necessary to introduce a new classification approach that is matched to these characteristics.

3. Insect Classification Using Scanning Electron Microphotographs

In this section, we propose insect classification using SEM images. In the proposed method, we prepare a dataset D that includes SEM images taken with various magnifications l_1, $l_2, \ldots, l_L$(L being the number of magnifications), and each image in the dataset D is labeled $1, 2, \cdots, T$(T being the number of species) in advance. Figure 2 shows an overview of the proposed method. In our method, we first generate the optimal training dataset for each magnification. Specifically, for each magnification l_i, we generate a training group including images of the target magnification l_i and images of its neighboring magnifications. In this procedure, we monitor classification accuracy by a leave-one-out scheme to find which images of the neighboring magnifications are useful for improving the classification accuracy of the target magnification l_i. Then it becomes feasible to construct the optimal classifier from the final grouped training dataset of each magnification. Furthermore, we calculate probabilistic outputs by using the optimal classifiers for each magnification.

In addition, we effectively use images of the same sample taken with different magnifications for improving the classification accuracy. Specifically, from the probabilistic outputs, the classification results are integrated on the basis of D-S evidence theory.

In the rest of this section, the details of our method are shown. In 3.1, we describe a method for visual feature extraction from SEM images and a feature selection method. In 3.2, we present a method for generating training datasets considering magnifications. In 3.3, we show derivation of probabilistic outputs using the support vector machine. The method for integration of results based on D-S evidence theory is presented in 3.4.

Note that insect classification is a multi-class classification problem, and two-class classification can solve the multi-class classification problem by a one-against-one method. Therefore, we first focus on two-class classification in 3.1 and 3.2 and describe our classification method based on two-class classification. Then we extend our method to multi-class classification by the one-against-one method in order to estimate multi-class probabilistic outputs and perform integration of results in 3.3 and 3.4.

3.1 Feature Extraction and Selection

In this subsection, visual features utilized in the proposed method are presented. Generally, we cannot obtain color information from SEM images since the images are grayscale images. Furthermore, images of each sample, i.e., each insect, are taken by an SEM from different angles or with different magnifications. Since SEM images have the above characteristics, the proposed method extracts SURF [20] and HOG features [25].

SURF is a robust feature for rotation and scaling. Since SURF is a local feature, we obtain a feature vector for each image by using BoF [21]. The main idea of BoF is representing images as collections of independent local patches and quantizing them as histogram vectors. In our method, the dimension of BOF is set to 500. HOG features are robust to local geometric and brightness changes as described in [25]. In order to extract HOG features from an image, we firstly divide the image into small spatial regions (blocks). Secondly, a histogram of each orientation in each region (cell) obtained by dividing the image into blocks is calculated. Finally, the histogram for each block is normalized and the HOG feature vector is obtained by concatenating the histograms of all blocks. In this paper, we use a block size of 3×3 cells and a cell size of $\frac{X}{6} \times \frac{Y}{4}$($X$ and Y being width and height of SEM images, respectively) pixels.

As described above, we extract feature vectors based on SURF and HOG features from SEM images. We denote this feature vector as $\boldsymbol{v} \in R^N$(N being a dimension of feature vector), and each element of $\boldsymbol{v}$ is represented as $v(j)(j = 1, 2, \cdots, N)$. Furthermore, the label $y \in \{1, -1\}$ represents species of insects for each image in the two-class problem. From N features, the proposed method selects $n(< N)$ features based on the mRMR algorithm [38]. Specifically, the method considers the redundancy of the features and the relevance of the features to the label and selects S_n, the set of n features, that maximizes the following equation:

$$\Phi(S_n) = D(S_n) - R(S_n), \tag{1}$$

where $D(S_n)$ is mutual information between $v(j) \in S_n$ and label y, which represents relevance; $R(S_n)$ is mutual information between two features $v(j)$ and $v(k) \in S_n$, which represents redundancy:

$$D(S_n) = \frac{1}{|S_n|} \sum_{v(j) \in S_n} I(v(j), y), \tag{2}$$

$$R(S_n) = \frac{1}{|S_n|^2} \sum_{v(j), v(k) \in S_n} I(v(j), v(k)), \tag{3}$$

where $I(\cdot, \cdot)$ represents the mutual information between two elements, and $|S_n|$ is the number of elements in S_n. Thus, the proposed method eliminates redundant features that cause performance degradation of the following insect classification procedures and enables the use of suitable features.

The above feature selection becomes feasible by using labeled training images. After performing feature selection, we define a new feature vector $\boldsymbol{x} \in R^n$ that contains n selected features. In this paper, we set n to 500.

3.2 Grouping Scheme for Making Training Dataset

In this subsection, a method for grouping images taken with different magnifications is presented. In the proposed method, we generate several datasets by grouping images taken with different magnifications. Then we calculate classification accuracy by using a leave-one-out scheme to obtain the dataset for which classification accuracy is the highest, where we use SVM[†] as the classifier to obtain the classification accuracy. Then the optimal training dataset that is optimal for each magnification can be obtained. The details are shown below.

From all of the images taken with the magnifications $l_1, l_2, \ldots, l_L$, we generate the training dataset for each magnification l_k by the following processes:

Step 1 Repeat the following procedures for all combinations of i, j $(1 \le i \le k \le j \le L)$.

- **(1-a)** A new dataset D_{ij} that consists of images with magnifications $l_i, \ldots, l_j$ is generated.
- **(1-b)** A training dataset and test dataset are obtained on the basis of leave-one-out scheme from the new dataset D_{ij}, where the test dataset is selected from only the target magnification l_k. Then we perform training of the classifier of SVM and verify its performance by using F-measure $F^k(i, j)$ shown as follows:

[†]The explanation of the SVM classifier used in our method is shown in the following subsection.

$$F^k(i,j) = \frac{2 \cdot r^k(i,j) \cdot p^k(i,j)}{r^k(i,j) + p^k(i,j)}, \tag{4}$$

$$r^k(i,j) = \frac{\mathrm{cor}^k(i,j)}{\mathrm{rel}^k(i,j)}, \tag{5}$$

$$p^k(i,j) = \frac{\mathrm{cor}^k(i,j)}{\mathrm{cla}^k(i,j)}, \tag{6}$$

where $\mathrm{cor}^k(i,j)$ is the number of correctly classified images, $\mathrm{rel}^k(i,j)$ is the number of relevant images, and $\mathrm{cla}^k(i,j)$ is the number of classified images.

Step 2 Find (i^k, j^k) that makes F-measure $F^k(i,j)$ the highest as follows:

$$(i^k, j^k) = \arg \max_{i \in \{1,\ldots,k\}, j \in \{k,\ldots,N\}} F^k(i,j). \tag{7}$$

Step 3 Define the dataset that consists of images with magnifications $l_{i^k}, \ldots, l_{j^k}$ as the training dataset for the target magnification l_k.

In this way, we can generate the training dataset that is optimal for each magnification l_k. Thus, it becomes feasible to improve the classification accuracy by considering magnifications of SEM images.

3.3 SVM-Based Probabilistic Output Calculation

In this subsection, a method for calculating a probabilistic output using SVM is presented. Given T classes of data, for any $\boldsymbol{x}$, the goal is to estimate

$$p_i = P(y = i|\boldsymbol{x}),\ i = 1, \cdots, T. \tag{8}$$

In the proposed method, we firstly select the training dataset of the magnifications $l_{i^k}, \ldots, l_{j^k}$ obtained as shown in Sect. 3.2 from the magnification l_k of the target image. Then we can prepare pairs $(\boldsymbol{x}_l^k, y_l^k)$ $(l = 1, \ldots, L^k)$ from the selected training dataset, where L^k is the number of images in the training dataset, $\boldsymbol{x}_l^k$ is the visual feature vector obtained from the l-th image in the training dataset, and y_l^k is the label. Furthermore, by using the training dataset, we calculate a separate hyperplane based on SVM [29] as follows:

$$f(\boldsymbol{x}) = \sum_{l=1}^{L^k} \alpha_l^k y_l^k K(\boldsymbol{x}, \boldsymbol{x}_l^k) + b^k, \tag{9}$$

where $K(\cdot,\cdot)$ is the kernel function for which the kernel is set as the linear kernel in this paper, α_l^k is a Lagrange multiplier, and b^k is a constant value. Since SVM is a two-class method, we utilize a one-against-one approach for multi-class classification. Thus, we first estimate pairwise class probabilities

$$r_{ij} \approx P(y = i|y = i \text{ or } j, \boldsymbol{x}) \tag{10}$$

using an improved implementation of Platt [39]. If $f(\boldsymbol{x})$ is the decision value at $\boldsymbol{x}$ obtained from Eq. (9), then we assume

$$r_{ij} \approx \frac{1}{1 + e^{Af(\boldsymbol{x})+B}}, \tag{11}$$

where A and B are estimated by minimizing the negative log likelihood of training data using their labels and decision values. After collecting all r_{ij} values, the following optimization problem is solved:

$$\min_{\boldsymbol{p}} \frac{1}{2} \sum_{i=1}^{T} \sum_{j=1, j \neq i}^{T} (r_{ji} p_i - r_{ij} p_j)^2,$$
$$\text{subject to } p_i \geq 0, \forall i,\ \sum_{i=1}^{T} p_i = 1, \tag{12}$$

where $\boldsymbol{p} = [p_1, \cdots, p_T]$ is a vector of the probabilities. By solving the optimization problem in (12), we can extend the two-class probability estimation to multi-class probability estimation. According to [40], we derive the probabilistic output p_k $(k = 1, 2, \cdots, T)$ from the estimated pairwise class probabilities r_{ij}. Specifically, the goal of Eq. (12) is to estimate $p_k = P(y = i|\boldsymbol{x})$ using all estimated pairwise class probabilities r_{ij}. According to the definitions of r_{ij} and p_k, the cost function of Eq. (12) can be regarded as the differences between two joint probabilities. Therefore, by minimizing these differences, the method in [40] tries to estimate p_k. Note that it is also shown in [40] that the optimization problem in Eq. (12) has a unique solution and can be solved by a simple linear system.

In this way, we can estimate probabilistic outputs using SVM. Then the proposed method can integrate probabilistic outputs based on D-S evidence theory described in the following subsection.

3.4 Integration of Probabilistic Outputs Based on D-S Evidence Theory

In this subsection, a method for integration of results for multiple images taken from the same sample, i.e., the same insect, is presented. For the target sample in the test data, we classify images $I_1, I_2, \ldots, I_E$(E being the number of images of the target sample taken with different magnifications). By the method presented in 3.3, we calculate the probabilistic output p_{jk} that image I_j is classified to species k. Specifically, the probabilistic output p_{jk} corresponds to p_k $(k = 1, 2, \cdots, T)$ of jth image I_j $(j = 1, 2, \cdots, E)$, where p_k can be obtained by solving Eq. (12). Then we integrate these outputs by D-S evidence theory. D-S evidence theory is a statistical-based data fusion classification algorithm. If Θ denotes the set of θ_k corresponding to T identifiable objects, let $\Theta = \{\theta_1, \cdots, \theta_T\}$ be a frame of discernment. The power set of Θ is the set containing all 2^T possible subsets of Θ, represented by $P(\Theta)$:

$$P(\Theta) = \{\Phi, \{\theta_1\}, \cdots, \{\theta_T\}, \{\theta_1, \theta_2\}, \cdots, \Theta\}, \tag{13}$$

where Φ denotes a null set. When the frame of discernment is determined, basic probability assignment (BPA) function $m(\cdot)$ mapping of the power set $P(\Theta)$ to $[0, 1]$ is defined by m :

$P(\Theta) \rightarrow [0, 1]$, and $m(\cdot)$ satisfies the following conditions:

$$\sum_{Q \in P(\Theta)} m(Q) = 1, \tag{14}$$

$$m(\Phi) = 0. \tag{15}$$

$m(Q)$ represents the proportion of all relevant and available evidence that supports the claim that a particular element of Θ belongs to the set Q but to no particular subset of Q. In this paper, the BPA function $m_j(k)$ is defined as

$$m_j(k) = p_{jk},\ j = 1, \cdots, E;\ k = 1, \cdots, T. \tag{16}$$

Evidence theory offers appropriate aggregation tools. Suppose $m_1(\cdot), m_2(\cdot), \cdots, m_E(\cdot)$ are E BPA functions formed on the basis of information obtained from E different information sources $I_1, I_2, \ldots, I_E$ in the same frame of discernment; according to Dempster's orthogonal rule [15], we have

$$\begin{aligned} m(Q) &= m_1(U_1) \oplus m_2(U_2) \oplus \cdots \oplus m_E(U_E) \\ &= \frac{\sum_{U_1 \cap \cdots \cap U_E = Q} m_1(U_1) \ldots m_E(U_E)}{1 - \sum_{U_1 \cap \cdots \cap U_E = \phi} m_1(U_1) \ldots m_E(U_E)}, \end{aligned} \tag{17}$$

where U_j is the subset of $P(\Theta)$ for information source I_i, and $\oplus$ is an operator that represents the integration of BPA functions. The decision function based on the maximal belief rule is

$$k^{opt} = \arg \max_{k=1,\cdots,T} \{\mathrm{Bel}(\{k\})\}, \tag{18}$$

$$\mathrm{Bel}(Q) = \sum_{C \subset Q} m(C),\ Q = 1, 2, \ldots, T. \tag{19}$$

In this process, we make the final decision k^{opt} of the target sample. D-S evidence theory is suitable for integration of some information sources. Therefore, the proposed method realizes improvement of classification accuracy by integration of probabilistic outputs obtained from the same sample.

4. Experimental Results

In this section, the effectiveness of the proposed method is verified from some experimental results using SEM images taken from insects. In order to verify the performance of the proposed method, we applied our method to the classification of family, superfamily and suborder. In addition, when we obtained classification results of some taxonomic ranks, performance improvement became feasible by integrating these results. Therefore, we show results of a method that integrates the results of multiple taxonomic ranks and verify its performance.

In 4.1, experimental results for each taxonomic rank are presented. In 4.2, we show the integration of results of multiple taxonomic ranks and its results. In 4.3, we show limitation of the proposed method and also discuss future work of our study.

4.1 Results for Each Taxonomic Rank

In this subsection, experimental results for each taxonomic rank are presented. We used images of insect foots taken by an SEM with magnifications from ×100 to ×10000. Table 1 shows the number of images for each magnification, and Fig. 3 shows the taxonomy of the dataset. Furthermore, in Table 2, we show the number of insects used for each family shown in Table 1. Each image is 8 bit grayscale and its size is 1280 × 840 pixels. In order to verify the effectiveness of the grouping scheme and the integration of results, we compared the performance of our method with the performances of the following comparative methods.

Comparative method 1
Only one classifier is constructed from all training dataset, and the test dataset is classified without considering magnifications.

Comparative method 2
A classifier is constructed from the training dataset obtained by the method presented in Sect. 3.2 for each magnification, and the test dataset is classified, but integration of results is no performed.

In addition, for comparing with other benchmarking methods, we performed classification by two classification methods (Local NBNN [33] and M-HMP [34]) that realize high accuracy for Caltech 101 and 256. Instead of using "Grouping scheme for making training dataset" (3.2) and "SVM-based probabilistic output calculation" (3.3), we used these two methods [33], [34] for verifying the effectiveness of our novel approach. The two previously reported methods, NBNN and M-HMP, provided high performance for Caltech 101 and 256 datasets, i.e., they realized successful classification for generic object recognition from limited number of training images. On the other hand, we realize the new method focusing on the unique characteristics of SEM images. Therefore, in order to confirm the effectiveness of the use of these unique characteristics, we performed the comparison with the previously reported methods that do not consider these characteristics.

Performance of each method was evaluated via 10-fold cross-validation. Therefore, from the total 2573 SEM images, we used about their 90% images as the training dataset and about 10% images as the test dataset, and iteratively performed the verification ten times with changing the test dataset to obtain their average performance. In order to evaluate the accuracy of the classification, we used Recall, Precision, F-measure and Accuracy shown in the following equations:

$$\mathrm{Recall} = \frac{\text{Num. of correctly classified images}}{\text{Num. of relevant images}}, \tag{20}$$

$$\mathrm{Precision} = \frac{\text{Num. of correctly classified images}}{\text{Num. of classified images}}, \tag{21}$$

$$\mathrm{F-measure} = \frac{2 \cdot \mathrm{Recall} \cdot \mathrm{Precision}}{\mathrm{Recall} + \mathrm{Precision}}, \tag{22}$$

Table 1 Number of images for each magnification.

	100	200	300	400	500	600	700	800	900	1000	1100	1200	1300	1400	1500
Hydrophilidae	79	47	19	9	26	3	2	6	4	21	1	3	1	0	1
Curculionidae	30	34	15	17	25	11	16	14	4	10	7	1	4	2	3
Cerambycidae	78	101	43	31	22	13	14	18	9	24	10	13	9	14	14
Chrysomelidae	50	56	61	35	22	7	18	15	13	11	9	10	7	12	13
Buprestidae	26	15	18	8	8	4	8	9	9	15	6	8	4	6	2
Carabidae	95	26	18	10	12	3	4	12	7	3	2	3	2	1	2
Gyrinidae	44	38	22	16	16	4	2	10	4	12	2	2	3	1	3
Dytiscidae	63	69	34	14	26	11	11	6	6	17	1	2	2	0	10
Cicindelidae	14	20	10	4	9	4	7	5	8	2	2	0	6	1	2
Reduviidae	23	27	12	15	12	6	1	4	2	2	0	0	2	2	3

	1600	1700	1800	1900	2000	2200	2300	2500	2700	3000	3300	3500	3700	4000	4300
Hydrophilidae	0	0	0	0	15	0	0	3	0	2	0	0	0	1	0
Curculionidae	2	0	4	2	4	3	3	1	1	0	1	1	0	2	0
Cerambycidae	4	7	4	9	6	5	3	9	6	7	2	3	2	2	4
Chrysomelidae	12	11	7	6	15	7	4	9	5	7	7	3	4	3	0
Buprestidae	3	1	1	5	4	3	1	3	1	3	1	4	1	3	0
Carabidae	0	0	3	2	2	2	1	1	1	0	1	2	0	1	0
Gyrinidae	1	0	4	1	4	1	2	0	1	2	0	2	0	0	0
Dytiscidae	1	0	1	1	11	3	1	5	0	5	3	5	1	7	0
Cicindelidae	1	1	0	0	0	0	3	0	1	0	0	0	0	0	0
Reduviidae	1	1	1	2	0	0	0	3	0	1	2	0	0	2	0

	4500	5000	5500	6000	7000	8000	8500	9000	9500	10000	Total
Hydrophilidae	0	3	0	0	1	0	0	1	0	1	249
Curculionidae	1	8	2	0	0	0	0	0	0	0	228
Cerambycidae	1	3	0	2	0	1	0	0	0	0	493
Chrysomelidae	3	1	0	1	0	0	1	0	0	0	445
Buprestidae	2	1	0	0	0	0	0	0	0	0	183
Carabidae	0	1	0	0	0	0	0	0	0	0	217
Gyrinidae	0	0	0	0	0	0	0	0	0	0	197
Dytiscidae	0	7	2	2	2	3	0	3	1	0	336
Cicindelidae	0	0	0	0	0	0	0	0	0	0	100
Reduviidae	0	1	0	0	0	0	0	0	0	0	125

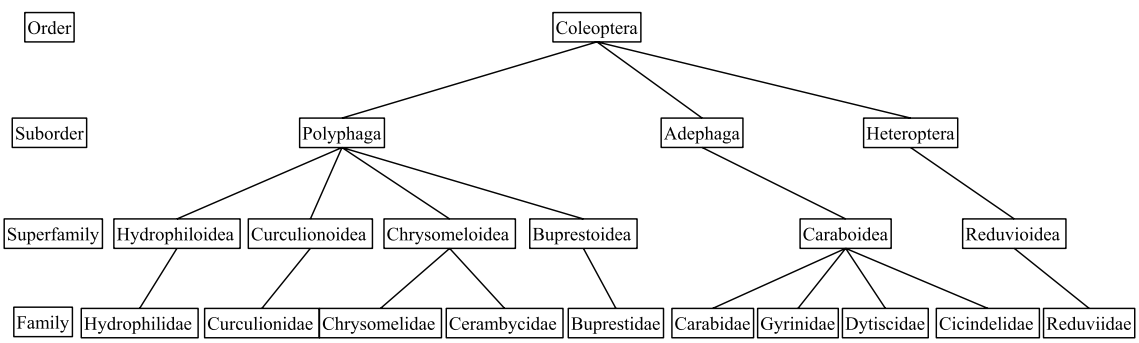


Fig. 3 Taxonomy of the dataset used in the experiment.

Table 2 Number of insects used for the verification.

Family	Num. of insects
Hydrophilidae	9
Curculionidae	9
Cerambycidae	8
Chrysomelidae	11
Buprestidae	3
Carabidae	12
Gyrinidae	6
Dytiscidae	14
Cicindelidae	2
Reduviidae	4

$$\text{Accuracy} = \frac{\text{Num. of correctly classified images}}{\text{Num. of images}}. \quad (23)$$

Tables 1–3 show the averages of Recall, Precision, F-measure and Accuracy. From these tables, we can see that the proposed method realizes a higher F-measure than those of the comparative methods. From the results obtained by using comparative method 1 and comparative method 2, the effectiveness of the grouping scheme can be verified. In addition, from the results obtained by using the proposed method and comparative method 2, we can see the integration of probabilistic outputs achieves improvement of classification accuracy. Furthermore, the classification results of M-HMP and Local NBNN are lower F-measure than that of the proposed method. These methods realize high accuracy for Caltech 101 and 256. However, since the target images are SEM images, performance degradation occurs due to the difference of visual features between different magnifications. From these experiments, we can confirm that the proposed method is suitable for classification of SEM images.

Table 3 Comparison of the performance of the proposed method and the performances of comparative methods for "family".

	Recall	Precision	F-measure	Accuracy
Proposed method	0.810	**0.756**	**0.782**	**0.778**
Comparative method 1	0.370	0.366	0.368	0.400
Comparative method 2	0.568	0.606	0.586	0.588
Local NBNN[33]	**0.872**	0.311	0.462	0.486
M-HMP[34]	0.654	0.643	0.654	0.654

Table 4 Comparison of the performance of the proposed method and the performances of comparative methods for "superfamily".

	Recall	Precision	F-measure	Accuracy
Proposed method	0.830	**0.781**	**0.801**	**0.805**
Comparative method 1	0.577	0.442	0.500	0.603
Comparative method 2	0.602	0.447	0.514	0.625
Local NBNN[33]	**0.872**	0.327	0.475	0.574
M-HMP[34]	0.653	0.641	0.462	0.485

Table 5 Comparison of the performance of the proposed method and the performances ot comparative methods for "suborder".

	Recall	Precision	F-measure	Accuracy
Proposed method	0.862	**0.803**	**0.834**	**0.851**
Comparative method 1	0.691	0.606	0.646	0.709
Comparative method 2	0.704	0.646	0.685	0.732
Local NBNN[33]	**0.897**	0.453	0.604	0.728
M-HMP[34]	0.692	0.724	0.713	0.767

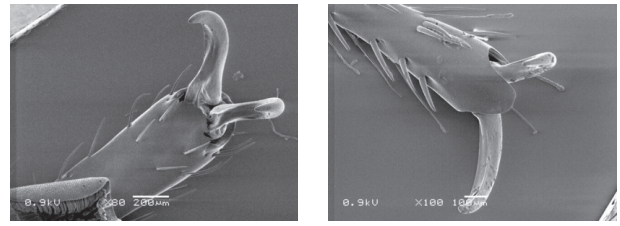
Table 6 Comparison of the performance of the method using only results for family and the method integrating results for family, superfamily and suborder.

	Recall	Precision	F-measure	Accuracy
Only family	0.810	0.756	0.782	0.778
Result integration	**0.832**	**0.781**	**0.814**	**0.803**

4.2 Integration of Results for Multiple Taxonomic Ranks

In this subsection, we show integration of results for multiple taxonomic ranks and evaluate the performance of the method by the results of an experiment. From the results shown in Tables 1–3, we can see that the classification performance of suborder and superfamily is better than that of family. Therefore, we can improve the performance of family by integration of the results for different taxonomic ranks, i.e., suborder and superfamily.

First, by the proposed method we classify SEM images for each taxonomic rank r = {suborder, superfamily, family} and calculate $\mathrm{Bel}(k^r)$ ($k^r = 1, \cdots, T^r$) in Eq. (19), where T^r is the number of elements belonging to taxonomic rank r. Specifically, $T^{\mathrm{suborder}} = 3$, $T^{\mathrm{superfamily}} = 6$, $T^{\mathrm{family}} = 10$ according to the taxonomy in Fig. 3. Then we integrate these results by adding $\mathrm{Bel}(k^{\mathrm{suborder}})$, $\mathrm{Bel}(k^{\mathrm{superfamily}})$ and $\mathrm{Bel}(k^{\mathrm{family}})$ according to the taxonomic rank, where $k^{\mathrm{family}} \in k^{\mathrm{superfamily}} \in k^{\mathrm{suborder}}$. For example, when we focus on Hydrophilidae (family), we add Bel(Hydrophilidae), Bel(Hydrophiloidea) and Bel(Polyphaga) according to the taxonomy in Fig. 3. Finally, we calculate $\mathrm{Bel}(\cdot)$ for all families and decide the family that becomes the highest value. Thus, it becomes feasible to improve the performance by integration of the results for different taxonomic ranks.



(a) (b)

Fig. 4 Examples of SEM images including different species that have common structures: (a) SEM images of Buprestidae at the magnification ×80, (b) SEM images of Cicindelidae at the magnification ×100.

Table 4 shows the averages of Recall, Precision, F-measure and Accuracy. From this table, the integration method realizes higher F-measure, and we can improve the performance of family by utilizing the results of suborder and superfamily. Thus, the effectiveness of the integration method was confirmed by this experiment.

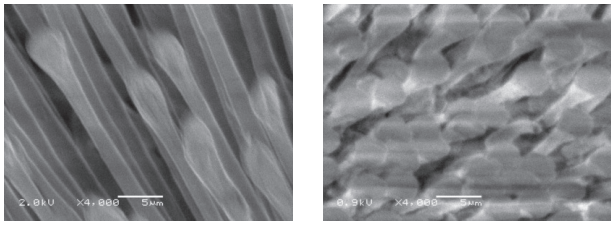
4.3 Limitation of Our Method and Its Future Work

In this subsection, we show the limitation of the proposed method and discuss its future work. We show some examples whose classification was difficult. Although the integration of the classification results based on D-S evidence theory is performed, the final classification performance becomes worse if the number of incorrect classification results becomes larger.

Generally, several different species have common structures which have similar functions as shown in Fig. 4. Thus, in such cases, since their visual features tend to become similar, accurate classification also becomes difficult. Furthermore, it becomes difficult to successfully classify SEM images of high magnification values due to the following two reasons. First, visual characteristics of SEM images of high magnification values tend to become similar since the contents included within these images become simple textures. Next, it is known that with increasing the magnification values, signal-to-noise-ratio (SNR) of SEM images tends to become lower [41]. This can be observed in Fig. 5. Therefore, in order to solve these problems, it is desirable to introduce new visual features which have high representation ability and robustness to the degradation of SEM images.

5. Conclusion

In this paper, we have proposed a method for insect classification using SEM images. In the proposed method, we construct an optimal classifier that enables classification for each magnification. In addition, we integrate the classification results for the same sample based on D-S evidence theory and realize effective insect classification. Finally, the



(a) (b)

Fig. 5 Examples of SEM images including some degradations due to the high magnifications: (a) SEM images of Curculionidae at the magnification ×4000, (b) SEM images of Reduviidae at the magnification ×4000.

effectiveness of the proposed method was evaluated by performing experiments using actual SEM images taken from insects.

Acknowledgement

This work was partly supported by JSPS KAKENHI Grant Number JP24120002 in Scientific Research on Innovative Areas "Innovative Materials Engineering Based on Biological Diversity".

References

[1] A.D. Chapman, Numbers of living species in Australia and the World, Department of the Environment, Water, Heritage and the Arts Canberra, 2009.
[2] P.D.N. Hebert, A. Cywinska, S.L. Ball, and J.R. deWaard, "Biological identifications through DNA barcodes," Proc. Royal Society B: Biological Sciences, vol.270, no.1512, pp.313–321, 2003.
[3] B. Bhushan, "Biomimetics: Lessons from nature-an overview," Philosophical Trans. Royal Society A: Mathematical, Physical and Engineering Sciences, vol.367, no.1893, pp.1445–1486, 2009.
[4] H. Zhang, Q. Huo, and W. Ding, "The application of AdaBoost-neural network in storedproduct insect classification," 2008 IEEE International Symposium on IT in Medicine and Education, pp.973–976, 2008.
[5] A. Lu, X. Hou, C.L. Liu, and X. Chen, "Insect species recognition using discriminative local soft coding," 21st International Conference on Pattern Recognition, pp.1221–1224, 2012.
[6] S.-G. Huang, X.-L. Li, M.-Q. Zhou, and G.-H. Geng, "SURF-based multi-scale resolution histogram for insect recognition," 2009 International Conference on Artificial Intelligence and Computational Intelligence, pp.445–448, 2009.
[7] X.-L. Li, S.-G. Huang, M.-Q. Zhou, and G.-H. Geng, "KNN-spectral regression LDA for insect recognition," 2009 First International Conference on Information Science and Engineering, pp.1315–1318, 2009.
[8] L.-Q. Zhu and Z. Zhang, "Auto-classification of insect images based on color histogram and GLCM," 2010 7th International Conference on Fuzzy Systems and Knowledge Discovery, pp.2589–2593, 2010.
[9] K. Venugoban and A. Ramanan, "Image classification of paddy field insect pests using gradient-based features," International Journal of Machine Learning and Computing, vol.4, no.1, pp.1–5, 2014.
[10] B. Shraddha, R. Charulata, M. Priyanka, and V.R. Pawar, "An automatic identification of agriculture pest insects and pesticide controlling," International Journal of Recent Research in Electrical and Electronics Engineering, vol.2, no.2, pp.21–28, 2015.
[11] M. Kearns and D. Ron, "Algorithmic stability and sanity-check bounds for leave-one-out cross-validation," Neural Comput., vol.11, no.6, pp.1427–1453, 1999.
[12] D. Zeng, J. Xu, and G. Xu, "Data fusion for traffic incident detector using D-S evidence theory with probabilistic SVMs," J. Computers, vol.3, no.10, pp.36–43, 2008.
[13] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Using the dempster-shafer method for the fusion of LIDAR data and multi-spectral images for building detection," Information Fusion, vol.6, no.4, pp.283–300, 2005.
[14] Z. Chang, X. Liao, Y. Liu, and W. Wang, "Research of decision fusion for multi-source remote-sensing satellite information based on SVMs and DS evidence theory," The 4th International Workshop on Advanced Computational Intelligence, pp.416–420, 2011.
[15] G. Shafer and A.F.M. Smith, "A mathematical theory of evidence.," Biometrics, vol.32, no.3, pp.703–704, 1976.
[16] A. Takahashi, T. Ogawa, and M. Haseyama, "Insect classification using scanning electron microphotographs considering magnifications," 2013 IEEE International Conference on Image Processing, pp.3269–3273, 2013.
[17] D.G. Lowe, "Object recognition from local scale-invariant features," Proc. 7th IEEE International Conference on Computer Vision, vol.2, pp.1150–1157, 1999.
[18] Y. Ke and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," Proc. 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2004, pp.506–513, 2004.
[19] A. Stein and M. Hebert, "Incorporating background invariance into feature-based object recognition," 2005 7th IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05), vol.1, pp.37–44, 2005.
[20] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up robust features (SURF)," Comput. Vis. Image. Und., vol.110, no.3, pp.346–359, 2008.
[21] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," Workshop on Statistical Learning in Computer Vision, pp.1–22, 2004.
[22] F. Jurie and B. Triggs, "Creating efficient codebooks for visual recognition," 10th IEEE International Conference on Computer Vision (ICCV'05), vol.1, pp.604–610, 2005.
[23] F. Perronnin, C. Dance, G. Csurka, and M. Bressan, "Adapted vocabularies for generic visual categorization," Computer Vision, ECCV 2006, Lecture Notes in Computer Science, vol.3954, pp.464–475, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
[24] J. van de Weijer and C. Schmid, "Coloring local feature extraction," Computer Vision, ECCV 2006, Lecture Notes in Computer Science, vol.3952, pp.334–348, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
[25] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), pp.886–893, 2005.
[26] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol.61, pp.85–117, 2015.
[27] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," Proc. ACM International Conference on Multimedia, MM'14, pp.675–678, 2014.
[28] H. Zhang, A.C. Berg, M. Maire, and J. Malik, "SVM-KNN: Discriminative nearest neighbor classification for visual category recognition," 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol.2, pp.2126–2136, 2006.
[29] C. Cortes and V. Vapnik, "Support-vector networks," Mach Learn., vol.20, no.3, pp.273–297, 1995.
[30] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Trans. Inf. Theory, vol.13, no.1, pp.21–27, 1967.
[31] A.S. Talaat, A.F. Atiya, S.A. Mokhtar, A. Al-Ani, and M. Fayek, "Multiclass penalized likelihood pattern classification algorithm," Neural Information Processing, Lecture Notes in Computer Science, vol.7665, pp.141–148, Springer Berlin Heidelberg, Berlin, Heidel-

berg, 2012.
[32] M.A. Kumar and M. Gopal, "Fast multiclass SVM classification using decision tree based one-against-all method," Neural Process. Lett., vol.32, no.3, pp.311–323, 2010.
[33] S. McCann and D.G. Lowe, "Local naive Bayes nearest neighbor for image classification," 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp.3650–3656, 2012.
[34] L. Bo, X. Ren, and D. Fox, "Multipath sparse coding using hierarchical matching pursuit," 2013 IEEE Conference on Computer Vision and Pattern Recognition, pp.660–667, 2013.
[35] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A.C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," Int. J. Comput. Vision., vol.115, no.3, pp.211–252, 2015.
[36] Y. Freund and R.E. Schapire, "Experiments with a new boosting algorithm," 13th International Conference on Machine Learning, vol.96, pp.148–156, 1996.
[37] J.D. Paola and R.A. Schowengerdt, "A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imagery," Int. J. Remote. Sens., vol.16, no.16, pp.3033–3058, 1995.
[38] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.8, pp.1226–1238, 2005.
[39] J.C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," Advances in Large Margin Classifiers, pp.61–74, 1999.
[40] T.F. Wu, C.J. Lin, and R.C. Weng, "Probability estimates for multi-class classification by pairwise coupling," J. Machine Learning Research, vol.5, pp.975–1005, 2004.
[41] N. Marturi, S. Dembélé, and N. Piat, "Performance evaluation of scanning electron microscopes using signal-to-noise ratio," 8th International Workshop on MicroFactories, pp.1–6, 2012.

Takahiro Ogawa received his B.S., M.S. and Ph.D. degrees in Electronics and Information Engineering from Hokkaido University, Japan in 2003, 2005 and 2007, respectively. He is currently an assistant professor in the Graduate School of Information Science and Technology, Hokkaido University. His research interests are multimedia signal processing and its applications. He has been an Associate Editor of ITE Transactions on Media Technology and Applications. He is a member of the IEEE, EURASIP, IEICE, and Institute of Image Information and Television Engineers (ITE).

Akihiro Takahashi received his B.S. and M.S. degrees in Electronics and Information Engineering from Hokkaido University, Japan in 2013 and 2015, respectively. He studied insect classification of scanning electron microscope (SEM) images using machine learning techniques.

Miki Haseyama received her B.S., M.S. and Ph.D. degrees in Electronics from Hokkaido University, Japan in 1986, 1988 and 1993, respectively. She joined the Graduate School of Information Science and Technology, Hokkaido University as an associate professor in 1994. She was a visiting associate professor of Washington University, USA from 1995 to 1996. She is currently a professor in the Graduate School of Information Science and Technology, Hokkaido University. Her research interests include image and video processing and its development into semantic analysis. She has been a Vice-President of the Institute of Image Information and Television Engineers, Japan (ITE), an Editor-in-Chief of ITE Transactions on Media Technology and Applications, a Director, International Coordination and Publicity of The Institute of Electronics, Information and Communication Engineers (IEICE). She is a member of the IEEE, IEICE, Institute of Image Information and Television Engineers (ITE) and Acoustical Society of Japan (ASJ).