

Misclassification correction in primary local recognition of component images of multichannel remote sensing data

Antti J. Niemistö^a, Vladimir V. Lukin^b, Alexander N. Dolia^b,
Olli P. Yli-Harja^a, Ilya Shmulevich^c

^aInstitute of Signal Processing, Tampere University of Technology, 33101, Tampere, Finland

^bDept. 504, National Aerospace University, 61070, Kharkov, Ukraine

^cUniversity of Texas, M.D. Anderson Cancer Center, TX 77030, USA

ABSTRACT

We propose an automatic classification procedure for multichannel remote sensing data. The method consists of several stages. An important stage is the correction of misclassifications based on the use of a nonlinear graph-based estimation technique recently introduced by us. The misclassification correction method is optimized by means of a training-based framework using genetic algorithms. It is shown that this provides a considerable improvement in classification accuracy. After primary local recognition and misclassification correction of all component images, an approach to further use the obtained data is considered. At this joint classification stage we introduce novel subclasses like “common homogeneous region”, “common edge”, “small sized object in one or two components”, etc. Numerical simulation data as well as real image processing results are presented to confirm the basic steps of remote sensing data classification and the efficiency of the proposed approach.

Keywords: remote sensing, image recognition, misclassification correction, multichannel images, training-based optimization, genetic algorithms

1. INTRODUCTION

One of the key problems in remote sensing is the design of automatic procedures for image classification¹. The existing remote sensing systems operating in different bands of electromagnetic waves (radio, infrared, optical) provide data arrays (images) that contain a great number of pixels. These images correspond to Earth terrains with areas of thousands of square kilometers. They offer the user the opportunity to carry out monitoring of various physical phenomena and of various scenes, i.e. to retrieve useful information by means of remote sensing². Obviously, for such applications it is not reasonable to expect that the required information can be reliably extracted by human experts by means of visual analysis of the obtained images. Even for highly qualified experts it is a difficult task to analyze such tremendous amounts of data.¹ The task is even more complicated if multichannel images are processed and classified.

By multichannel remote sensing data we mean both images jointly registered by radar imaging subsystems that are characterized by different operation frequencies (wavelengths) or polarizations³ and multispectral images obtained in the optical band⁴. These kinds of remote sensing systems are considered to be the most useful, since the use of several spectral or polarization channels provides an opportunity to essentially increase the information content of remote sensing data and the reliability of the required information retrieval^{3,4}. On the other hand, the multichannel mode of remote sensing leads to new challenges in image processing⁵. One of them is the question of how to classify these images automatically.

One important property of multichannel images is that their component images are rather similar (correlated) in general, but at the same time they can be considerably different for some particular areas^{2,6}. This phenomenon can be easily explained by the fact that different types of sensed objects are characterized by different backscattering or irradiation properties dependent upon object type and mechanisms of the sensing signal interaction with the considered object type. The combination of these similarities and differences is the basis of the multichannel approach to remote sensing and information retrieval from the obtained remote sensing data. One typical example is the

Further author information:

A.N.: E-mail: antti.niemisto@tut.fi

V.L.: E-mail: lukin@ai.kharkov.com

detection of thin films of oil or other pollutions on sea surface in multichannel radar images. In the short wavelength radar component images (for example, in K_a -band) these polluted areas can be detected while in the images formed by radars with larger wavelength this information can be practically absent⁷. This means that the information about edges and small sized objects and their locations in different component images registered jointly can be very useful for practical applications.

In automatic classification of multichannel images misclassifications often occur in the neighborhoods of edges and small sized objects⁸. This is natural since neighborhood information is commonly used for classification. That is why in the neighborhood of edges and small sized objects the confusing situations can occur with a much higher probability than in homogeneous image regions where the neighboring pixels belong to the same class^{4,9}. Further, it is often advantageous to use different techniques for classification of homogeneous regions and small sized objects in multichannel images. For homogeneous image regions the basic features used for classification are the intensities of different components or their ratios⁹ while for classification of the small sized objects it is reasonable to additionally use the spatial features like object size, shape, etc.

Taking all this into account we can conclude that automatic classification of multichannel remote sensing data is a challenging task. The presence of noise in real images makes this task even more complicated. One reasonable solution is the division of the automatic classification procedure into separate stages and the design of efficient methods for getting the appropriate results at each stage. In this paper, we propose to divide the multichannel image classification procedure into two stages. The first one is the separate classification of component images, and the second stage is the so called “joint” classification of the multichannel image based on the separate classification. The first stage is divided into two substages, i.e. primary recognition and secondary misclassification correction which are discussed in Sects. 2 and 3, respectively. The simulation results of the first stage of classification of multichannel images are described in Sect. 4 in connection with the application of a training-based optimization framework for the misclassification correction method. The joint classification stage is discussed in Sect. 5.

2. PRIMARY LOCAL RECOGNITION OF COMPONENT IMAGES

The first stage of classification of multichannel images is the separate classification of each component image. Our approach consists of two steps (substages), namely primary local recognition (PLR) of component images and misclassification correction of component images (MCCI). As the result of PLR each pixel of each component image gets classified into one of the following subclasses:¹⁰

- homogeneous region (H);
- edge neighborhood between two homogeneous regions having different intensity means (E);
- spike neighborhood, i.e. there is a spike in the scanning window, but not in the central pixel (NS);
- spike in the central pixel of the scanning window – this corresponds to impulsive noise, i.e. outliers (S);
- central pixel belongs to a small sized or prolonged object – a small sized object is characterized by compactness and connectedness of pixels belonging to it as well as by homogeneity of values (O);
- neighborhood of a small sized object, i.e. there is a small sized or prolonged object in the scanning window, but the central pixel does not belong to it (NO).

The reasons for such PLR are the following. Firstly, if a spike is detected, it is desirable to replace the corresponding pixel value by some other value. This is done taking into account the pixel values in the neighborhood of the spike using, for example, some nonlinear filtering algorithm such as the center weighted median filter¹¹. This is needed since the presence of spikes can dramatically decrease the reliability of image classification. Another goal is to localize edges, because then it will become easier to make the classification in their neighborhoods since one can then refer the corresponding pixels to one of the neighboring homogeneous regions that form the edge. Finally, the detection of small sized objects permits the analysis of their spatial features, and thus the exploitation of additional parameters for a more reliable classification.

PLR using the above subclasses can be done rather easily if the component images are noise-free, and a recognition method for such images has been described by Niemistö et al¹². The method is based on the calculation of the local

histogram in a scanning window fashion and uses simple decision rules to distinguish between subclasses. It is used in Sect. 4 to produce a target image for training-based optimization.

In the noisy case, the recognition task is more challenging. Let us first review the image/noise models that we use. The first one assumes that the image is corrupted by mixed Gaussian multiplicative and impulsive noise. Then, a noisy component image can be expressed as

$$x_{ij} = \begin{cases} I_{ij}\mu_{ij}, & \text{with probability } 1 - P_I \\ A_{ij}, & \text{with probability } P_I \end{cases}, \quad (1)$$

where x_{ij} is a pixel of the noisy image at i th row and j th column, I_{ij} is the respective true value of the pixel and μ_{ij} denotes multiplicative noise with the mean equal to unity and relative variance equal to σ_μ^2 . P_I is the probability of the occurrence of an impulse and A_{ij} is the corresponding value of the pixel corrupted by an impulse. This model is typical in, for example, radar imaging systems, which are in turn typical sources of input data for image recognition systems.¹³

The second model corresponds to the case of an image corrupted by mixed Gaussian additive and impulsive noise. The model for a component image is

$$x_{ij} = \begin{cases} I_{ij} + \chi_{ij}, & \text{with probability } 1 - P_I \\ A_{ij}, & \text{with probability } P_I \end{cases}, \quad (2)$$

where χ_{ij} denotes additive noise with zero mean and variance equal to σ_χ^2 . The model (1) can be reduced to the model (2) using a homomorphic transformation, and therefore we only need to consider the latter model. In practice this means that we can use the same PLR system irrespective of whether we are dealing with multiplicative or additive noise.¹⁴

A method for PLR of noisy images has been proposed by Lukin et al¹⁰ and Dolia et al^{14,15}. Naturally, the method can be applied separately to the components of a multichannel image. The recognition is done locally using a scanning window of size 5×5 . The size is large enough for reliable recognition of the central pixel, and at the same time it is small enough to avoid ambiguity in recognition of complex situations. For a given position of the scanning window six local statistical parameters are evaluated. These parameters are then used as input data for the classifier (PLR system), which can be either an expert system¹⁰ or a three-layer radial basis neural network¹⁵. The local statistical parameters are chosen such that their behavior is rather different for the subclasses that we consider. Then, by means of a ‘‘joint analysis’’ of these parameters, the classifier can assign the most probable hypothesis (subclass) to the central pixel in the scanning window.

The components of our first test image, a three-channel 256×256 image which we call ‘‘Mosaic’’, are shown in Fig. 1. In the first component image, there is a homogeneous region in the lower part and in the middle horizontal part there are small sized objects having different shapes and different numbers of pixels belonging to them. Their contrasts with respect to the background vary as well. In the upper part there are homogeneous regions inside and between mosaic type objects with different shapes, orientations and contrasts. In the second and third component images the homogeneous region is larger and extends to the top of the image on the right hand side. In this way we can simulate the properties of real remote sensing images that are described in the Introduction.

The components of the ‘‘Mosaic’’ image are corrupted by multiplicative Gaussian noise with relative variances $\sigma_\mu^2 = 0.003$, $\sigma_\mu^2 = 0.006$ and $\sigma_\mu^2 = 0.02$, respectively. The values are chosen such that they are close to the relative noise variances that are observed in practice in multichannel side-looking aperture radar (SLAR) images. In these images, the components are formed by different subsystems with different characteristics. For example, we can have one component that is formed by a SLAR with operation wavelength 8 mm and relative noise variance of about 0.005, while another component could be formed by a SLAR with a wavelength of 3 cm and relative noise variance of about 0.012.⁵

In the ‘‘Mosaic’’ image the noise is purely multiplicative, i.e. $P_I = 0$. However, Dolia et al¹⁴ have proposed a PLR method that can produce a good recognition result even in the presence of impulsive noise, and therefore the absence of impulses is not important. The results of PLR of the noisy component images are shown in Fig. 2, and we see that there are quite a few misclassifications in all of the component images. The elimination of these misclassifications is a specific task for secondary processing of the PLR results.

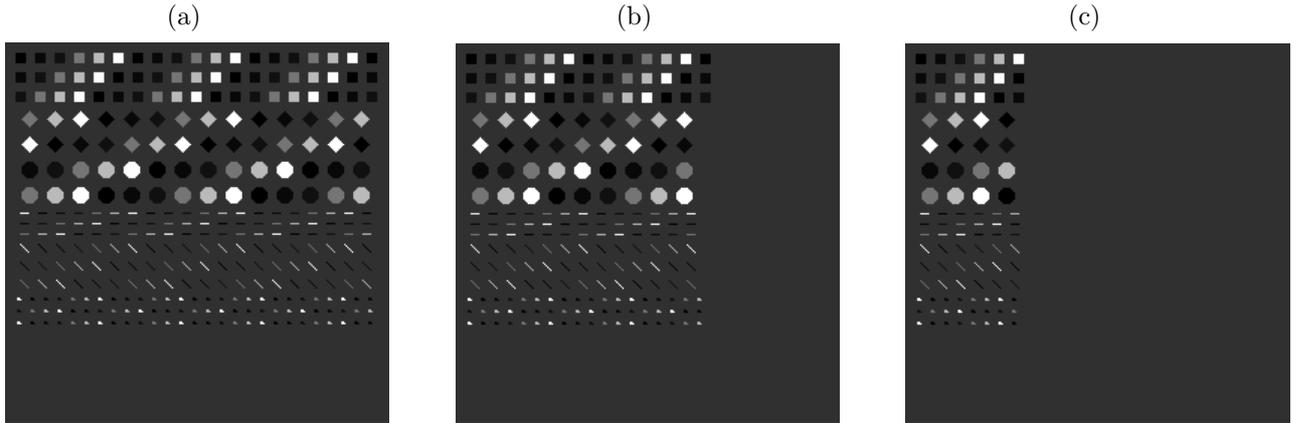


Figure 1. The three components of the noise-free “Mosaic” test image.

3. MISCLASSIFICATION CORRECTION OF COMPONENT IMAGES

The following requirements are defined for the correction of misclassifications after primary classification¹⁶:

- R1. edge (E) pixels and small sized object (O) pixels should be preserved if they are spatially grouped and they are surrounded by NO and NS pixels;
- R2. a spike (S) should be preserved if it is surrounded by NS pixels;
- R3. isolated pixels of H, E, S and O should be removed;
- R4. all pixels with initial classification NS or NO should be removed.

Since all pixels with initial classification NS or NO should be removed, these neighborhood subclasses are considered supplementary subclasses. In fact, their primary purpose is to assist in misclassification correction, and as can be seen by looking at requirements R1 and R3, the distinction between correctly recognized and misclassified E and O pixels is made primarily on the basis of whether they are surrounded by supplementary subclasses or not. Further, in this paper the spike subclass (S) is considered as supplementary, since we use our noise models such that the probability of an impulse is zero ($P_I = 0$). Generally, this assumption is not necessary as explained in Sect. 2. The subclasses H, E and O are considered as basic subclasses.

Let us consider the graph-based estimation technique (GET) for correction of misclassifications that was introduced by Yli-Harja et al¹⁷ and further developed by Lukin et al¹⁶ as well as Niemistö et al¹². The method is based on a complete directed weighted graph $G(V, E)$, where V is the vertex set and E is the edge set. Let the weight function be $w : V \times V \mapsto \mathbf{R}$. It is well known that in the case of real numbers, the L_p -norm estimate of $X = \{x_1, x_2, \dots, x_N\}$, $x_i \in \mathbf{R}$ is the value β that minimizes $\sum_{i=1}^N |x_i - \beta|^p$. Suppose now that we have some set of samples $A = \{v_1, v_2, \dots, v_N\}$, $v_i \in V$ of graph G . Then we can define

$$\text{graph-}p(A) = \arg \min_{\beta \in A} \sum_{i=1}^N w(\beta, v_i)^p \quad (3)$$

to be the graph-based “ L_p -norm” estimate. The value $p = 1$ corresponds to the graph-based “median” and $p = 2$ to the graph-based “mean”. Note that the estimate is necessarily one of the vertices in A . We also lift the requirement of a metric space. This means that the weights do not have to be symmetric, i.e. it is possible that $w(v_i, v_j) \neq w(v_j, v_i)$ for $i \neq j$, and that they do not have to obey the triangle inequality. Moreover, as the following example illustrates, vertices may be repeated in A , i.e. it is possible that $v_i = v_j$ for $i \neq j$, and thus A is in fact a multiset.

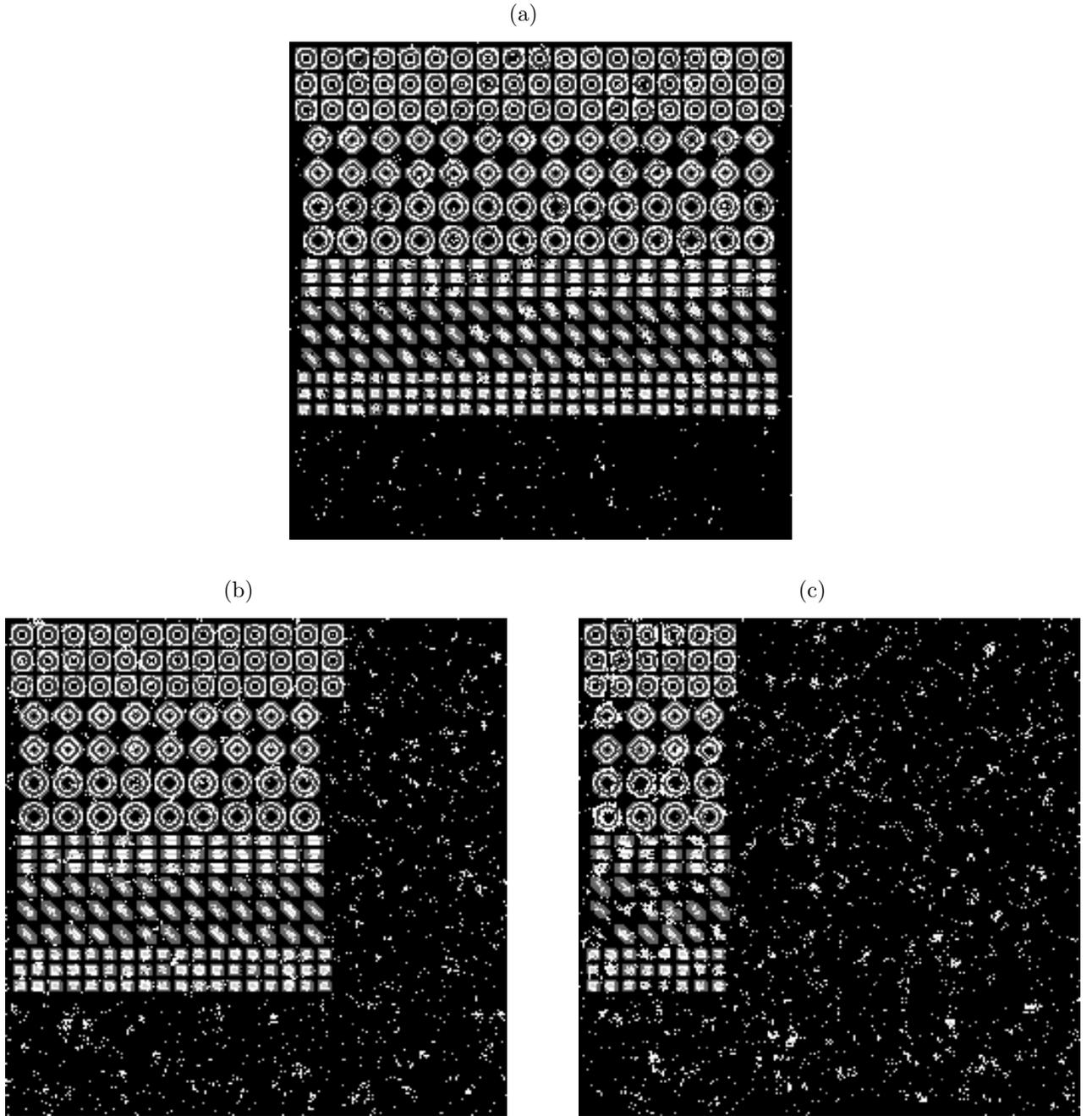


Figure 2. The primary local recognition results of the components of the noisy “Mosaic” image with subclasses H, E, NS, S, O and NO shown by gradations of gray from black to white.

A simple modification to (3) can be made in order to ensure that the output always belongs to the basic subclasses. Let A_b be the set of samples in the window that represent the basic subclasses and let A_s be the set of samples in the window that represent the supplementary subclasses. Then $A = A_b \cup A_s$ and $A_b \cap A_s = \emptyset$ and we can define¹²

$$\text{graph-}p(A, A_b) = \arg \min_{\beta \in A_b} \sum_{i=1}^N w(\beta, v_i)^p. \quad (4)$$

This modification guarantees the removal of supplementary subclasses, since the output will by definition always belong to one of the basic subclasses. Then, similarly to the running median and mean filters, we can define a scanning window processing operation based on (4). Here we use $\text{graph-1}(A, A_b)$ as our processing operator, i.e. we calculate the graph based “median” for each position of the scanning window.

Next, the size of the scanning window must be chosen. Since a 5×5 window is used in PLR, it is natural to use a 5×5 window here as well. This is because, ideally, a 5×5 neighborhood of an E or O pixel contains those supplementary subclass (NS and NO) pixels that are meant to assist in correction of misclassifications. A larger window would include a lot of H pixels, and correction of misclassifications would be more difficult. A smaller window, in turn, would not include all the information contained in the supplementary subclasses, because some of them would be left out of the window.

Finally, it is also possible to weight the different samples in the scanning window differently. The approach in which weighting is done by duplication of samples is chosen here, in a manner similar to weighted median filtering. For simplicity, only the central sample is weighted with the weight $k = 10$, although any other symmetric choice of weights is possible and reasonable. The weighting of the central sample increases the probability of the central sample remaining as the output. This is desired, since it decreases the probability of a pixel that has been correctly recognized in PLR to be misclassified in secondary misclassification correction. However, at the same time the probability that a misclassified pixel remains misclassified increases, and therefore center weighting is a compromise between detail preservation and misclassification correction.

4. TRAINING-BASED OPTIMIZATION

The utilization of a training-based optimization framework for finding the optimal graph for GET has been proposed by Niemistö et al¹². In this section, we review the optimization phase and discuss the results of secondary misclassification correction using the optimal graph. This finalizes the first stage of the classification procedure for multichannel images.

In training-based optimization using a genetic algorithm we have a pair (S, f) , where S is a set of configurations (solution space) and $f : S \mapsto \mathbf{R}$ is a fitness function. The fitness function, f , is the optimization criterion to be maximized, and the problem is to find a configuration $s_0 \in S$ such that

$$f(s_0) \geq f(s) \quad \forall s \in S. \quad (5)$$

The value of the fitness function is calculated iteratively for candidates, s , in a population. In each iteration the genetic algorithm applies simple genetic operations to the population in order to increase the fitness of its individuals. In principle, the algorithm converges when the optimal configuration, s_0 , is found as one of the individuals in the population.¹⁸ In our case, each candidate, s , is a weight function $w : V \times V \mapsto \mathbf{R}$ of the candidate graph. The weight function is represented by an adjacency matrix whose entries give the weights between the vertices of the graph.

Obviously, the choice of fitness function is crucial. In our case, S is the set of all possible adjacency matrices, and the only input parameter of the fitness function is a candidate, $M \in S$, for the optimal adjacency matrix (the initial candidate is chosen randomly). The real number to be maximized is then $f(M)$. This number is calculated as follows:

1. process the (partially misclassified) source image using the method described in Sect. 3 using the current candidate M ;
2. calculate the number, N , of pixels in the source image that agree with the corresponding pixels of the target image;
3. return N^2 .

In the last step, any monotonically increasing function of N could be returned. Of course, the simplest choice would be to just return N . However, genetic algorithms are quite sensitive to the choice of the fitness function, and it turns out that in this case using N^2 is a good choice.

The genetic optimization algorithm that we use is taken from Ref. 18. We choose to use 3 bits to represent each weight of the adjacency matrix, which means that each of them can take 8 different values. The bounds for

the weights can be chosen arbitrarily, and since natural numbers are easier to manipulate than real numbers, we choose $w(v_i, v_j) \in [0, 7]$ for all pairs of i and j . Making these choices, the solution space for the genetic algorithm is small enough in order for it to converge and at the same time it is large enough for the genetic algorithm to find a satisfactory optimization result. The number of weights of the adjacency matrix to be optimized is 18 (see below).

Let us use the first component of the partially misclassified “Mosaic” image, shown in Fig. 2 (a), as the source image for training-based optimization. The target image, obtained by PLR of the noise-free “Mosaic” image, is shown in Fig. 3 (a). The training phase results in the optimal adjacency matrix

$$M_{opt} = \begin{matrix} & \text{H} & \text{E} & \text{NS} & \text{S} & \text{O} & \text{NO} \\ \begin{matrix} 0 \\ 6 \\ - \\ - \\ 5 \\ - \end{matrix} & \begin{bmatrix} 6 & 0 & 3 & 5 & 3 \\ 3 & 4 & 4 & 2 & 4 \\ - & - & - & - & - \\ - & - & - & - & - \\ 7 & 1 & 3 & 0 & 6 \\ - & - & - & - & - \end{bmatrix} & \end{matrix}. \quad (6)$$

Because a pixel representing the subclasses S, NS or NO cannot be an output of (4), the weights in the third, fourth and sixth row do not have an effect on the output of (4), and therefore do not undergo optimization. To emphasize this, those weights are labeled as “-” in the matrix above.

Let us now apply GET with the optimal adjacency matrix (6) to the components of the “Mosaic” image. We show that although the matrix was trained for the component image with the smallest relative noise variance ($\sigma_\mu^2 = 0.003$), it can be applied to the other component images with higher noise levels as well. The component images after secondary misclassification correction are shown in Fig. 3 (b-d). The first observation is that in all the images no misclassifications remain in the homogeneous regions (H). Moreover, the subclasses E and O are preserved well, although in the third component some of the small sized objects are missing entirely. The results of misclassification correction are shown numerically in Table 1. The size of the component images is 252×252 .

Table 1. The numerical results of secondary misclassification correction of the components of the “Mosaic” image.

Component	First	Second	Third
Misclassifications (1)	1070	1426	1178
Supplementary	19494	13899	6745
Misclassifications (2)	822	688	430
Corrected	471	815	982
Misclassified	65	10	126
Misclassified (suppl)	158	68	108

The first and second rows of Table 1 give, for each component image, the number of misclassifications and the number of pixels representing the supplementary subclasses after PLR, respectively. The third row gives the number of misclassifications after MCCI. It can be seen that for each component image the number of misclassifications is reduced by the application of GET. Moreover, an appropriately large number of misclassifications is corrected and an appropriately small number of new misclassifications are introduced in each case as can be seen on the fourth and fifth rows, respectively. Finally, the last line indicates that only a small percentage of pixels representing the supplementary subclasses are replaced by the wrong subclass.

Naturally, the misclassification correction result is good for the component image for which GET is trained, namely the first component image. This is true even though the number of misclassifications is largest for the first component, because the number of edges (E) and small sized objects (O), which are harder to classify than homogeneous regions (H), is largest for the first component. Further, the relative amount of corrected E and S pixels is comparable for the second and third component images. Namely, the number of corrected E pixels is 148, 18 and 37 for the first, second and third component images, respectively, and the same numbers for O pixels are 57, 43 and 19, respectively. Thus, a reasonable misclassification correction result is obtained for all the components of the “Mosaic” image.

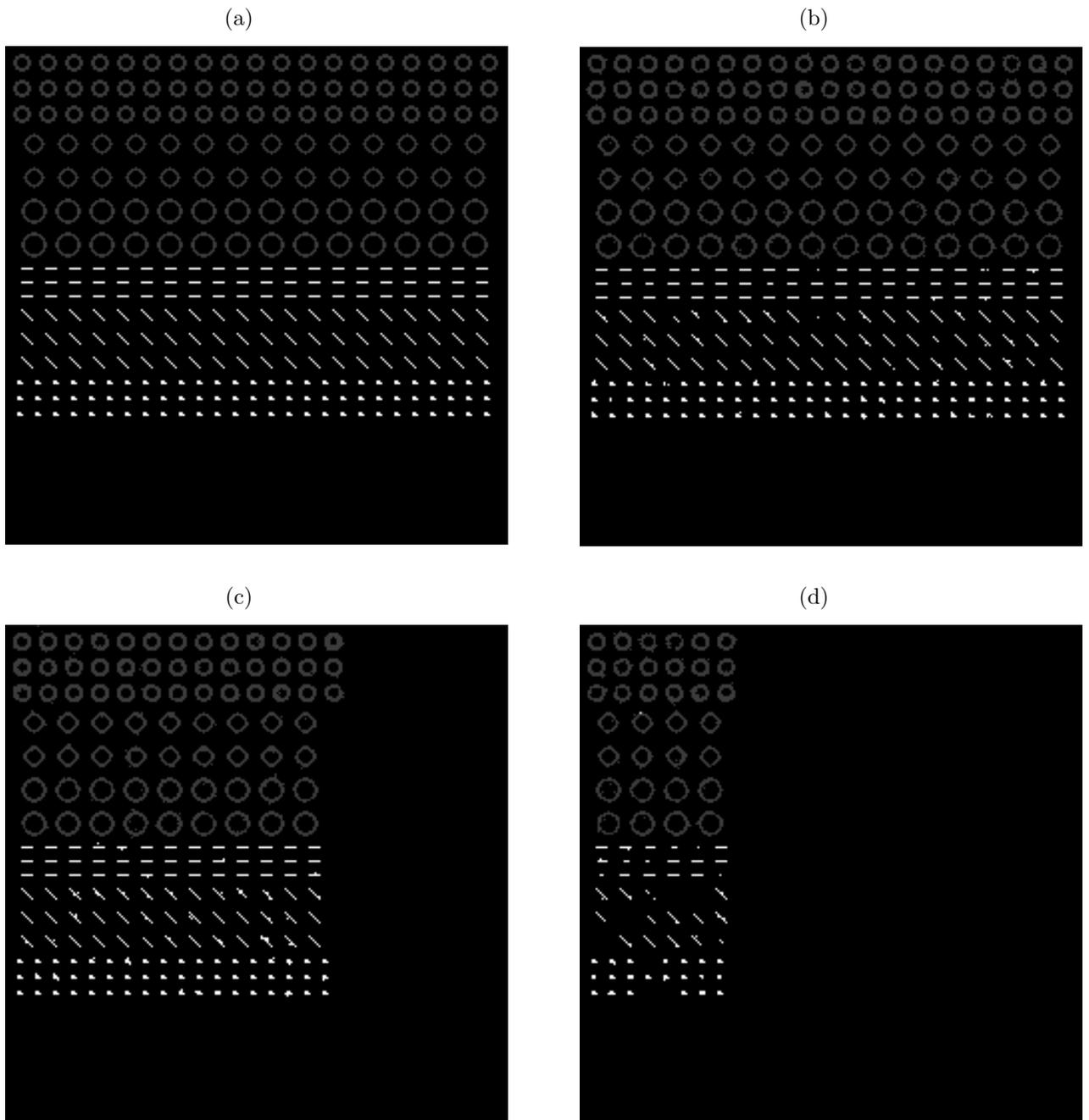


Figure 3. (a) The “Mosaic” target image for training-based optimization; (b), (c), (d) the secondary misclassification correction results of the components of the “Mosaic” image. The subclasses are H (black), E (gray) and O (white).

The above approach to classification of component images is also applied to the real 8-bit SLAR image presented in Fig. 4 (a). The multiplicative noise in this image has the relative variance approximately equal to 0.006. The image corresponds to an agricultural region in Ukraine and contains bare soil areas, forest areas (mainly in the upper part), a sunflower field (in the lower right corner) and some buildings and gardens around them (a small village). The small sized (prolonged) objects are gravel roads and, in some places, tree strips protecting the fields from wind. The image was obtained in October, i.e. at a time when some fields had already been ploughed. The results of PLR are given in Fig. 4 (b). It is seen that the contrast edges and small sized objects are detected well. Their neighborhoods

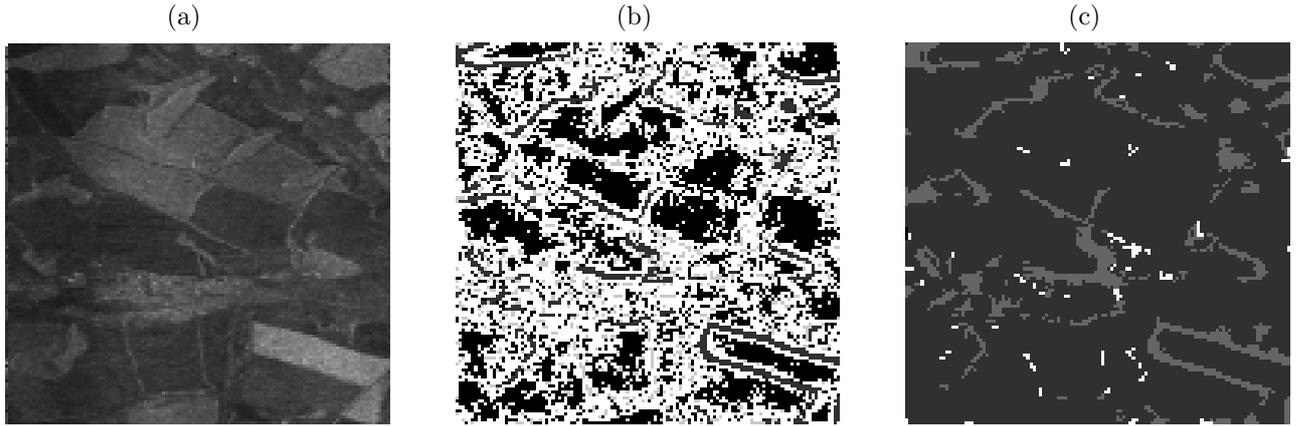


Figure 4. (a) A SLAR image of an agricultural region in Ukraine; (b) the primary local recognition results with subclasses H (black), E (dark gray), O (light gray) and NO (white); (c) the misclassification correction results with subclasses H (black), E (gray) and O (white).

are also recognized well and are indicated by the supplementary subclass NO (white pixels). Homogeneous regions like forest areas and the sunflower field are mainly indicated by H (black pixels). However, for many other areas, mainly corresponding to bare soil fields and the village there is a mixture of subclasses with many pixels classified as NO. This is because these areas are, in fact, not homogeneous (the bare soil field is characterized by spatially varying erosion state resulting in backscattered signal intensity variation in rather wide limits⁹).

We also apply the misclassification correction procedure to the SLAR image. The adjacency matrix (6) is used, and the result is shown in Fig. 4 (c). In addition to misclassification correction using GET, further misclassification correction has been achieved by means of removing isolated E and O pixels by applying a simple majority filter to these pixels. The H pixels are shown by black, E pixels are gray and O pixels are indicated by white. This image is “clean” and much more clearly shows the homogeneous regions (H), edges (E) and small sized objects (O). However, the classification results obtained for the real SLAR image show that in the future it is also worth introducing a PLR subclass called “textured areas”.

Finally, the results of PLR and MCCI in the case of two- or three-channel radar image processing can be represented in a way that seems useful for remote sensing data analysis. Multichannel radar images are often represented as color images where each component radar image is associated with the corresponding color component in the RGB space¹⁹. However, here we propose to use a color representation of component image classification results. As the first step, the subclasses are associated with constant values. Specifically, for pixels recognized as H the value 0 is assigned, for E pixels we assign the value 128, and for O pixels the value 255. This operation is performed for all component images. Then, in the case of three-channel radar remote sensing, those multichannel image pixels that are recognized as H for all components will get the color black, pixels recognized as E for all components will be shown by a gray color and pixels recognized as O in all components will be indicated by the color white. If a small sized object is recognized in only one component then it will get a color close to the corresponding color component. The same applies to an edge recognized in only one component. By means of this color representation it is possible to make the visual detection and localization of edges and small sized objects which are present in only one component image easier. Thus, some particular tasks like detection of oil pollutions on sea surface mentioned in the Introduction can be more easily solved at the stage of preliminary analysis and classification of multichannel remote sensing data.

5. JOINT CLASSIFICATION OF MULTICHANNEL IMAGES

The third stage of the classification process is the “joint” classification and misclassification correction of the components of the multichannel image. We use the term joint classification to distinguish the classification procedure described in this section from the PLR and MCCI stages described above. The idea in joint classification is to assign a single subclass for each pixel of the multichannel image, i.e. for an image with n channels (components) the choice of subclass depends on n pixels. In the case of three channels we propose the following five subclasses:

- Common homogeneous region, i.e. the classification in all three components is H (CH);
- Common edge, i.e. the classification in all three components is E (CE);
- Common small sized objet, i.e. the classification in all three components is O (CO);
- Edge in one or two components and homogeneous region in the other component(s) (PE);
- Small sized object in one or two components and homogeneous region in the other component(s) (PO).

The choice of the set of subclasses such as above depends on the data to be classified. In our case, classification of multichannel (multifrequency) radar remote sensing images is a potential application. In these images, situations justifying the introduction of the subclasses PE and PO often occur. As explained in the Introduction this is due to the fact that radars operating in different frequencies recognize different image features, and it may happen that in some channels we have edges corresponding to, for instance, the margins of an area with a thin oil film, and in other channels these edges are absent or not detectable.

The joint classification of multichannel images can be done pixel by pixel by examining the component images and choosing the appropriate subclass. The component images for which secondary correction of misclassifications has been performed are used here, i.e. in the case of the “Mosaic” image the component images shown in Fig. 3 (b-d) are used. In some cases it may happen that none of the proposed subclasses is appropriate. For example, it may happen that in one of the components there is an edge (E) pixel and in the other two there is a small sized object (O) pixel, or vice versa. Naturally, these kinds of situations are misclassifications by definition. However, in practice these misclassifications do not occur very often, and e.g. for the entire “Mosaic” image this happens only once. Therefore, we classify these situations as CH, since this is the most probable classification for real remote sensing data. This decision also simplifies the training-based optimization task, since the number of parameters to be optimized can be reduced to 25 when we do not introduce new subclasses for this kind of exceptional situations (see below).

The above considerations are also the reason behind the selection of the components of the “Mosaic” image, i.e. the selection was made such that all of the above subclasses appear in the image after joint classification. The result of joint classification of the “Mosaic” image using the above subclasses is shown in Fig. 5 (a). We see immediately that the classifier performs well on the right hand side of the image. However, on the left hand side there are more misclassifications. This is due to the fact that in the third component image, which has the highest noise level and thus also the greatest number of misclassifications, there are edges (E) and small sized objects (O) only on the left.

Like in separate classification of component images, we can apply GET for the task of correction of misclassifications. However, since we do not have supplementary subclasses in this case, we use a smaller scanning window than in Sect. 3, i.e. we use a 3×3 window. The weight of the central sample is $k = 10$ in order to achieve good detail preservation. Then, training-based optimization can be applied to find the optimal graph for this case. We use the same training method as in Sect. 4 and obtain the optimal adjacency matrix

$$M_{joint} = \begin{matrix} & \text{CH} & \text{CE} & \text{CO} & \text{PE} & \text{PO} \\ \begin{bmatrix} 0 & 5 & 4 & 6 & 7 \\ 4 & 1 & 5 & 1 & 4 \\ 1 & 2 & 2 & 6 & 3 \\ 6 & 7 & 1 & 2 & 3 \\ 1 & 4 & 7 & 7 & 5 \end{bmatrix} & & & & & \end{matrix} \quad (7)$$

The “Mosaic” image after joint misclassification correction (and removal of isolated pixels) using the above matrix is shown in Fig. 5 (b). We can see that many of the misclassifications have been corrected, in particular on the left hand side of the image. Specifically, the number of misclassifications before misclassification correction is 1424 and after the correction it is 1182. In more detail, the classification of 246 pixels have been corrected and only 4 correctly classified pixels have been replaced by the wrong subclass. Moreover, there are no discontinuities in the edges (CE or PE) and also the small sized objects (CO and PO) are preserved well. Thus, it is clear that GET can produce significant correction of misclassifications also in the joint classification stage.

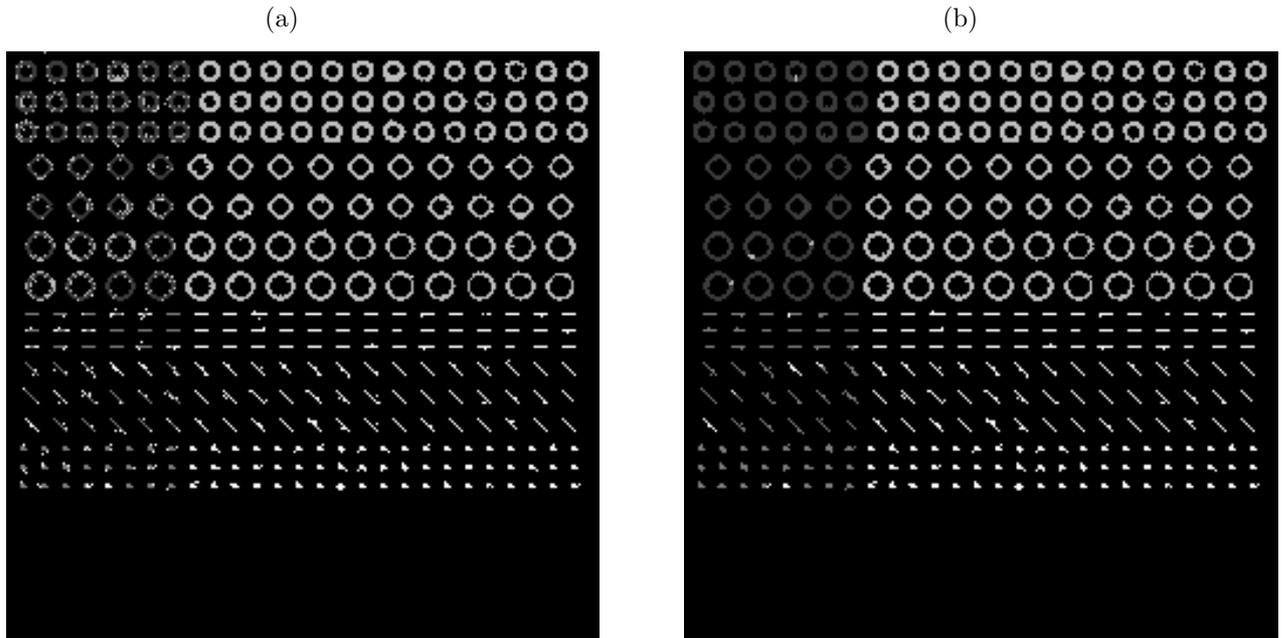


Figure 5. (a) The joint classification result of the “Mosaic” image; (b) the “Mosaic” image after joint misclassification correction. The subclasses are CH, CE, CO, PE and PO, and are shown by gradations of gray from black to white.

6. CONCLUSION

An automatic classification procedure for multichannel remote sensing data is proposed. The first stage of the procedure is based on primary local recognition of component images, and the second stage is based on a joint analysis of the primary recognition results. At the second stage some novel subclasses relating to multichannel images are introduced. Further, after both of the stages classification accuracy is improved by correcting misclassifications using a graph-based estimation technique, GET. This is done in a training-based optimization framework using genetic algorithms. It is shown that correction of misclassifications is crucial in classification of multichannel images. Moreover, numerical simulation data are presented to confirm the basic steps of classification of multichannel remote sensing data and the efficiency of the proposed approach.

Further, the first stage of the classification procedure, namely primary local recognition and misclassification correction, is also applied to a real SLAR image. It is shown that the classification result is reasonable. Future work will concentrate on applying the classification method on real multichannel remote sensing images with the possible introduction of a novel subclass corresponding to textured areas.

REFERENCES

1. R. Congalton, K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, CRC Press, Boca Raton, 1999.
2. J. Richards, *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag, Berlin, 1994.
3. W. Keydel, “SAR techniques and technology, its present state of the art with respect to user’s requirements,” in *Proc. of the European Conf. on Synthetic Aperture Radar*, pp. 19–24, Konigswinter, Germany, Mar 1996.
4. D. Landgrebe, “Information extraction principles and methods for multispectral and hyperspectral image data,” in *Information Processing for Remote Sensing, chapter one*, C. H. Ghen, ed., World Scientific Publishing, 1999.
5. V. Lukin, J. Astola, V. Melnik, A. Kurekin, A. Zelensky, G. Kulemin, N. Ponomarenko, A. Dolia, J. Parkkinen, “Data fusion and processing for airborne multichannel system of radar remote sensing: methodology, stages, and algorithms,” in *Sensor Fusion Architectures, Algorithms, and Applications IV*, B. V. Dasarathy, ed., *Proc. SPIE* **4051**, pp. 215–226, 2000.

6. K.-H. Szekiedla, "Satellite Monitoring of the Earth", Wiley, New York, 1989.
7. G. Jolly, W. Alpers, V. Barale, F. Cauneau, G. van der Graaf, A. Mangin, J. Redondo, O. Rud, H. Snaith, "Preliminary results from the clean sea projects: pollution monitoring from space," in *Future Trends in Remote Sensing*, P. Gudmandsen, ed., Balkema, Rotterdam, 1998.
8. B. Solaiman, L. Pierce, F. Ulaby, "Multisensor data fusion using fuzzy concepts: application to land-cover classification using ERS-1/JERS-1 SAR composites," in *IEEE Trans. on Geoscience and Remote Sensing* **37**, pp. 1316–1326, May 1999.
9. A. Kurekin, V. Lukin, A. Zelensky, O. Tsymbal, G. Kulemin, E. Engman, "Processing multichannel radar images by modified vector sigma filter for soil erosion degree determination," in *Remote Sensing for Earth Science, Ocean, and Sea Ice Applications*, G. Cecchi, ed., *Proc. SPIE* **3868**, pp. 412–423, 1999.
10. V. Lukin, N. Ponomarenko, J. Astola, K. Saarinen, "Algorithms of nonlinear adaptive filtering using fragment recognition by expert system," in *Nonlinear Image Processing VII*, E. R. Dougherty, ed., *Proc. SPIE* **2662**, pp. 179–190, 1996.
11. S.-J. Ko, Y. H. Lee, "Center weighted median filters and their applications to image enhancement," in *IEEE Trans. on Circuits and Systems* **38**, pp. 984–993, Sep 1991.
12. A. Niemistö, V. Lukin, I. Shmulevich, O. Yli-Harja, A. Dolia, "A training-based optimization framework for misclassification correction," in *Proc. of the 12th Scandinavian Conf. on Image Analysis*, pp. 691–698, Bergen, Norway, Jun 2001.
13. V. Melnik, *Nonlinear Locally Adaptive Techniques for Image Filtering and Restoration in Mixed Noise Environments*, Ph. D. Thesis, Publications 286, Tampere University of Technology, Tampere, Finland, 2000.
14. A. Dolia, V. Lukin, A. Zelensky, J. Astola, C. Anagnostopoulos, "Neural networks for local recognition of images with mixed noise," in *Applications of Artificial Neural Networks in Image Processing VI*, N. M. Nasrabadi, ed., *Proc. SPIE* **4305**, pp. 108–118, 2001.
15. A. Dolia, A. Burian, V. Lukin, C. Rusu, A. Kurekin, A. Zelensky, "Neural network application for primary local recognition and nonlinear adaptive filtering of images," in *Proc. of the 6th IEEE Int. Conf. on Electronics, Circuits and Systems*, pp. 847–850, Pafos, Cyprus, Sep 1999.
16. V. Lukin, I. Shmulevich, O. Yli-Harja, A. Dolia, "Correction of misclassifications in primary local image recognition using a nonlinear graph-based estimation technique," in *Image and Signal Processing for Remote Sensing VI*, S. B. Serpico, ed., *Proc. SPIE* **4170**, pp. 251–259, 2000.
17. O. Yli-Harja, I. Shmulevich, K. Lemström, "Graph-based smoothing of class data with applications in musical key finding," *Proceedings of the IEEE–EURASIP Workshop on Nonlinear Signal and Image Processing*, pp. 311–315, Antalya, Turkey, Jun 1999.
18. D. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading, Massachusetts, 1989.
19. K. Saarinen, *On Filtering, Analysis, and Display of Digital Color and Radar Images*, Ph. D. Thesis, Publications 282, Tampere University of Technology, Tampere, Finland, 1999.