Fusion of Multisensor Data Based on Different Multidimensional Distributions

Aivars Lorencs[1], Ints Mednieks[1], Martins Pukitis[1], Juris Sinica-Sinavskis[1]
[1]Institute of Electronics and Computer Science,
Dzerbenes iela 14, LV-1006 Riga, Latvia
mednieks@edi.lv

***Abstract*—The paper addresses problems related to classification of images obtained by various types of remote sensing devices. Development and use of Bayes type land cover classifiers based on multidimensional Gaussian, Dirichlet and gamma distributions is analysed and compared on the basis of sample data from RGB and hyperspectral thermal sensing devices with unequal spatial resolution. Approaches to data fusion for design of the combined classifiers are presented including the cases where different families of multidimensional distributions are used to model the sensor data and classifiers are designed using combinations of their probability density functions. The best classification results are obtained when the fusion of data from both images is used together with classification based on all three considered distributions combined together.**

***Index Terms*—Data fusion; image classification; multidimensional distribution; remote sensing.**

I. INTRODUCTION

Development of accurate and fast methods for classification of remotely sensed image pixels to predefined classes remains one of the major tasks in Earth observation from satellites and planes. To increase the amount of available information and raise the classification accuracy, combination of data from different sensors is often used thus adding complexity to the task. In the paper [1], Bayes classification approach to this task was presented, based on the assumption that each class can be interpreted as a sample realisation from the multidimensional universe with Gaussian distribution. However, it was noticed that this assumption may not fit well to data for several classes. Therefore, in this study an attempt was made to use different multidimensional distributions [2], [3], develop related classifiers and data fusion approaches, and perform comparative analysis. As in [1], it was done on the basis of the "subset" images of the data set "grss_dfc_2014" presented for the 2014 IEEE GRSS Data Fusion Contest (DFC) [4].

II. INITIAL DEFINITIONS AND ASSUMPTIONS

Input data sources were 2 images (I, I') of the same land surface area captured in two different regions of spectrum

and with different spatial resolution. I is a visible image made by airborne high resolution RGB camera during one overflight of the area; I' is the image of the same area taken by LWIR hyperspectral device in 84 spectral bands with wavelengths from 7.8 μm to 11.5 μm. Pixels of the RGB image are represented by vectors $\mathbf{x} = (x_1, x_2, x_3)^T$ where x_i is the sensed light intensity of the corresponding color. Pixels of the other image are represented by vectors $\mathbf{u} = (u_1, u_2, ..., u_{84})^T$, where u_i is the sensed radiation intensity in the spectral band with a central wavelength λ_i. Further we will call this image TI (thermal infrared) image. Spatial resolution of RGB images is $a \approx 0.2 \times 0.2$ m while that of the TI image is $b \approx 1 \times 1$ m, and b is not a multiple of a.

Classification of pixels into 7 categories (roads, trees, red roof, grey roof, concrete roof, vegetation and bare soil) was considered using the ground truth areas of all categories made available in the image I by the organizers of the contest. These pixels were used to design the intended classifiers and therefore they are elements of the design set. In this study, the designed classifiers were tested using the same set of pixels in the same way as it was done in [1] and with the same justification. As I' is the image of the same scene, ground truth areas depicted in I can be transformed into the corresponding areas related to the same categories in image I' and pixels within these areas form the design set for the TI image.

Let us denote by S the design set of pixels in image I and by S' the design set of pixels in image I'. Subsets of these sets related to the category k, $k = \overline{1,7}$ will be denoted correspondingly by S_k and S'_k.

From the TI image, 8 spectral bands (4, 14, 26, 36, 47, 57, 69, and 78) out of 84 were chosen in the same way as in [1] so that pixels of this image were represented by 8-dimensional vectors $\mathbf{y} = (y_1, y_2, ..., y_8)^T$. The problem of selection of the more informative bands from the hyperspectral image TI is not considered in this paper.

On the basis of this research lies the hypothetic postulate that the intensity vectors $\mathbf{x}$ resp. $\mathbf{y}$ of a category can be qualified as sample realizations from a multidimensional universe $\mathbf{X}_k$ resp. $\mathbf{Y}_k$. This postulate is supplemented by the plausible assumption that all probability distributions of

$\mathbf{X}_k$ resp. $\mathbf{Y}_k$, $k = \overline{1,7}$ belong to the same family of distributions.

All pixel classifiers designed within this research are constructed as Bayes type classifiers based on corresponding probability distributions. As there is no information available about the prior probabilities of categories, they are assumed to be equal.

The goal of this study was to assess the influence of different probability distributions and their combinations to the accuracy of classifiers constructed on their basis both for separate classification of single images and using data fusion.

III. FAMILIES OF MULTIDIMENSIONAL PROBABILITY DISTRIBUTIONS

From the families of multidimensional probability distributions described in literature [2], [3] we chose the Gaussian, Dirichlet and gamma distributions for analysis because their probability density functions are easy to use in analytical expressions. Generalized Gaussian distribution [3] was initially considered as well but it was noticed that, in most cases, commonly used Gaussian distribution, which is a specific case of the generalized Gaussian distribution, provided the best classification accuracy or differed from the best result insignificantly. Using Gaussian distribution is additionally driven by the fact that various authors consider it to be the adequate model for optical images [5], [6]. Of course, it cannot be considered an ideal model as all components of the intensity vectors $\mathbf{x}$ and $\mathbf{y}$ are non-negative values. For the same reason, Dirichlet and gamma distributions deserve attention as well. Obviously, neither Dirichlet nor gamma distribution can be referred to pixel intensity vectors directly. Dirichlet distribution can describe only vectors $\mathbf{x} = (x_1, x_2, ..., x_n)^T$, for which $x_i \geq 0$, $i = \overline{1,n}$, and $\sum_{i=1}^{n} x_i = 1$. As to chosen case of gamma distribution, it is necessary that $x_i \geq 0$, and $x_{i+1} \geq x_i$. There are several ways to overcome these "obstacles". To apply the Dirichlet distribution, vectors $\mathbf{x}$ can be transformed to vectors $\breve{\mathbf{x}} = (x_1/c, x_2/c, ..., x_n/c)^T$, where $c = \sum_{i=1}^{n} x_i$. The only deficiency of this transformation is the fact that two vectors with proportional components will be transformed to the same vector. We can avoid this situation by choosing the positive real value c' as a divider, which is greater than all possible values of c. It however leads to the necessity of introduction of an additional variable $x_{n+1} = 1 - \sum_{i=1}^{n} x_i$.

For the chosen case of gamma distribution, it is also necessary to perform a transformation in order to meet conditions set up for this distribution. Here we need to transform vector $\mathbf{x}$ into $\mathbf{x}^* = (x_1^*, x_2^*, ..., x_n^*)^T$, where $x_1^* = x_1$, $x_2^* = x_1 + x_2$, ..., $x_n^* = \sum_{i=1}^{n} x_i$.

The described transformations are performed both for RGB and TI images. As the probability density functions of considered distributions depend on the particular parameter values, construction of the classifiers should be started from the estimation of these parameters for each category of pixels.

IV. ESTIMATION OF DISTRIBUTION PARAMETERS

To use the n-dimensional Gaussian distribution as a model for the pixel vectors of category k, it is necessary to calculate the estimates of parameters $\mathbf{\mu}_k$ and $\mathbf{\Sigma}_k$ of the probability density function

$$f_{\mathbf{X}_k}(\mathbf{x}) = (2\pi)^{-n/2} |\mathbf{\Sigma}_k|^{-1/2} \times$$
$$\times \exp\left(-1/2(\mathbf{x} - \mathbf{\mu}_k)^T \mathbf{\Sigma}_k^{-1}(\mathbf{x} - \mathbf{\mu}_k)\right), \quad (1)$$

where $\mathbf{\mu}_k = E(\mathbf{X}_k)$ is the mean value and $\mathbf{\Sigma}_k = E\left[(\mathbf{X}_k - \mathbf{\mu}_k)(\mathbf{X}_k - \mathbf{\mu}_k)^T\right]$ is the covariance matrix.

As the probability density function for the Dirichlet distribution is equal to

$$f_{\mathbf{X}}(x_1, x_2, ..., x_{n-1}) = \frac{\Gamma(v)}{\prod_{i=1}^{n} \Gamma(v_i)} \prod_{i=1}^{n-1} x_i^{v_i - 1} \left(1 - \sum_{i=1}^{n-1} x_i\right)^{v_n - 1}, \quad (2)$$

where $v = \sum_{i=1}^{n} v_i$, $\sum_{i=1}^{n-1} x_i \leq 1$, $x_i \geq 0$, $v_i > 0$, then it is necessary to develop formulas for the estimation of parameters v_i.

It is proven (see [2]) that the mean value of marginal distribution X_i of the random vector $\mathbf{X} = (X_1, X_2, ..., X_{n-1})^T$ is equal to $E(X_i) = v_i/v$ but variance $\text{var}(X_i) = v_i(v - v_i)/(v^2(v+1))$. It means that $v_i = vE(X_i)$. In this case, $\text{var}(X_i) = E(X_i) \times (1 - E(X_i))/(v+1)$, from where we obtain that $v = (E(X_i)(1 - E(X_i)) - \text{var}(X_i))/\text{var}(X_i)$. Hence

$$v_i = \left((E(X_i))^2(1 - E(X_i)) - E(X_i)\text{var}(X_i)\right)/\text{var}(X_i). \quad (3)$$

By substituting $E(X_i)$ and $\text{var}(X_i)$ with their statistical estimates obtained from the pixel intensity vectors belonging to S_k, we get estimates $\hat{v}_{i,k}$ of v_i and a plausible analytical expression for the Dirichlet probability density function of the pixel intensity vectors of category k

$$\breve{f}_{\mathbf{X}_k}(\breve{x}_1, \breve{x}_2) = \frac{\Gamma\left(\sum_{i=1}^{3} \hat{v}_{i,k}\right)}{\prod_{i=1}^{3} \Gamma(\hat{v}_{i,k})} \prod_{i=1}^{3} \breve{x}_i^{\hat{v}_{i,k} - 1}, \quad (4)$$

where $\breve{x}_3 = 1 - \breve{x}_1 - \breve{x}_2$, $\hat{v}_{3,k} = \frac{1}{2}[\hat{v}' + \hat{v}''](1 - E(X_1) - E(X_2))$,

$\hat{v}' = ((E(X_1))(1 - E(X_1)) - \text{var}(X_1))/\text{var}(X_1)$,

$$\hat{v}'' = \left(\left(E(X_2) \right) \left(1 - E(X_2) \right) - \mathrm{var}(X_2) \right) / \mathrm{var}(X_2).$$

Estimates $\hat{v}_{i,k}$ should be soundly based on the transformed intensity vectors $\breve{\mathbf{x}}$. If the vector $\mathbf{x}$ is substituted by vector $\hat{\mathbf{x}} = \left(\hat{x}_1, \hat{x}_2, \hat{x}_3 \right)^T$, and the additional variable $\hat{x}_4 = 1 - \sum_{i=1}^{3} \hat{x}_i$ is introduced, then we obtain the plausible density function for the Dirichlet model

$$f_{\mathbf{X}_k}(\hat{x}_1, \hat{x}_2, \hat{x}_3) = \frac{\Gamma\left(\sum_{i=1}^{4} \hat{v}_{i,k} \right)}{\prod_{i=1}^{4} \Gamma(\hat{v}_{i,k})} \prod_{i=1}^{4} \hat{x}_i^{\hat{v}_{i,k} - 1}. \qquad (5)$$

If we consider the gamma distribution to be the model of probability distribution of pixel intensity vector for category k, then it should be taken into account that, in n-dimensional case, this probability density function is analytically described by the equality

$$f_{\mathbf{X}}(\mathbf{x}^*) = \frac{\left(x_1^* - \gamma_1 \right)^{\alpha_1 - 1}}{\beta^{\alpha_n^*} \prod_{i=1}^{n} \Gamma(\alpha_i)} \times$$

$$\times \prod_{i=2}^{n} \left(x_i^* - x_{i-1}^* - \gamma_i \right)^{\alpha_i - 1} \exp\left(-\frac{x_n^* - \sum_{i=1}^{n} \gamma_i}{\beta} \right), \qquad (6)$$

where $\alpha_i > 0$, $\alpha_n^* = \sum_{i=1}^{n} \alpha_i$, $\beta > 0$, $x_i^* \geq \gamma_i \geq 0$. By considering intensity vectors $\mathbf{x}$ of pixels of category k and corresponding transformed vectors $\mathbf{x}^*$, we can assume that all $\gamma_i = 0$.

It is proven (see [2]) that $E(X_i) = \alpha_i^* \beta$ and $\mathrm{var}(X_i) = \beta^2 \alpha_i^*$, where $\alpha_i^* = \sum_{j=1}^{i} \alpha_j$. Hence $\beta = \mathrm{var}(X_i) / E(X_i)$, $\alpha_1 = \left(E(X_1) \right)^2 / \mathrm{var}(X_1)$, $\alpha_2 = \left(\beta \alpha_2^* - \beta \alpha_1 \right) / \beta = \left(E(X_2) - E(X_1) \right) E(X_1) / \mathrm{var}(X_1)$, …, $\alpha_n = \left(\beta \alpha_n^* - \beta \alpha_{n-1}^* \right) / \beta = \left(E(X_n) - E(X_{n-1}) \right) E(X_1) / \mathrm{var}(X_1)$.

By substituting $E(X_i)$ and $\mathrm{var}(X_i)$ with their statistical estimates obtained from the pixel intensity vectors belonging to S_k, we get estimates $\hat{\alpha}_k$ and $\hat{\beta}_k$. As a result, we obtain the probability density function of the possible model for the gamma distribution

$$\breve{f}_{\mathbf{X}_k}(\mathbf{x}^*) = \frac{\left(x_1^* \right)^{\hat{\alpha}_{1,k} - 1}}{\hat{\beta}^{\alpha_{3,k}^*} \prod_{i=1}^{3} \Gamma(\hat{\alpha}_{i,k})} \times$$

$$\times \left(x_2^* - x_1^* \right)^{\hat{\alpha}_{2,k} - 1} \left(x_3^* - x_2^* \right)^{\hat{\alpha}_{3,k} - 1} \exp\left(-\frac{x_3^*}{\hat{\beta}_k} \right). \qquad (7)$$

All parameter estimation formulas and analytical expressions of the probability density functions, with necessary formal changes, remain in force also for the TI image.

V. CLASSIFICATION OF SEPARATE IMAGES

At the first stage, classifiers K_i, ($i = \overline{1,3}$ is the index of used probability distribution; K_1 is the classifier based on Gaussian distribution; K_2 is the classifier based on Dirichlet distribution; K_3 is the classifier based on gamma distribution) targeting differentiation of pixels within the set S between 7 categories are constructed, using only information from image I. In the same way, classifiers K_i' targeting differentiation of pixels within the set S' between 7 categories are constructed using only information from image I'.

Let us denote by $f_{k,i}$ the probability density function defined using the parameter estimates of the corresponding distribution, calculated from the pixel intensity vectors $\mathbf{x}$ of category k taken from the design set S. In an analogous way, we denote by $g_{k,i}$ the probability density function defined using the parameter estimates of the corresponding distribution, calculated from the pixel intensity vectors $\mathbf{y}$ of category k taken from the design set S'.

The separate Bayes classifier K_i is defined by the condition that it classifies a pixel r with the intensity vector $\mathbf{x}$ from the set S as a pixel of category k if and only if the following inequality is met

$$f_{k,i}(\mathbf{x}) / f_{j,i}(\mathbf{x}) > 1, \qquad (8)$$

for all $j \neq k$. This inequality can be replaced by the equivalent inequality

$$\ln f_{k,i}(\mathbf{x}) - \ln f_{j,i}(\mathbf{x}) > 0. \qquad (9)$$

In an analogous way, the classifier K_i' is defined: K_i' classifies pixel r' from the set S' with intensity vector $\mathbf{y}$ as a pixel of category k if and only if

$$g_{k,i}(\mathbf{y}) / g_{j,i}(\mathbf{y}) > 1, \qquad (10)$$

for all $j \neq k$. Inequality (10) is replaced by the equivalent inequality

$$\ln g_{k,i}(\mathbf{y}) - \ln g_{j,i}(\mathbf{y}) > 0. \qquad (11)$$

VI. CLASSIFICATION USING DATA FUSION

Combination of classifiers using data from single image is possible in several ways. Our experiments were focused on 3 ways of designing classifier $U_{i,j}$ for classification of pixels from data set S using fusion of classifiers K_i and K_j':

Option 1. $U_{i,j}$ classifies pixel r from RGB image with

intensity vector **x** as a pixel of category k if and only if

$$\ln f_{k,i}(\mathbf{x}) - \ln f_{s,i}(\mathbf{x}) + \ln g_{k,j}(\mathbf{y}^a) - \ln g_{s,j}(\mathbf{y}^a) > 0, \quad (12)$$

for all $s = \overline{1,7}$, $s \neq k$, $\mathbf{y}^a$ is the intensity vector of the pixel r' from the TI image, associated with pixel r as it corresponds to the same area, namely includes the area related to pixel r. To eliminate possible errors related to edges of defined ground truth areas where the associated pixels r' can include places of different category, initial ground truth areas were cropped by applying morphologic erosion.

Option 2. $U_{i,j}$ classifies pixel r from RGB image with intensity vector **x** as a pixel of category k if and only if

$$\tfrac{2}{3}\left(\ln f_{k,i}(\mathbf{x}) - \ln f_{s,i}(\mathbf{x})\right) +$$
$$+ \tfrac{1}{3}\left(\ln g_{k,j}(\mathbf{y}^a) - \ln g_{s,j}(\mathbf{y}^a)\right) > 0, \quad (13)$$

for all $s = \overline{1,7}$, $s \neq k$. This option provides the possibility to use different weights for data from images RGB and TI (2/3 and 1/3 is just an example).

Option 3. Classify each image separately using separate classifiers K_i and K_j'; if both results are the same, accept it; if they differ:

– Sort the values of probability density functions obtained for each image for all categories in descending order;
– Choose the decision of the classifier for which the difference between the first (highest) and the second values of the probability density function is larger.

Combination options of classifiers described above can be applied without significant statement changes to classification of individual images as well. Thereby we can define *e.g.* classifiers $V_{i,j}$ and $V_{i,j}'$, where $V_{i,j}$ is a combination result of the RGB image classifiers K_i and K_j, but $V_{i,j}'$ is a combination result of the TI image classifiers K_i' and K_j'. Obviously, such combined classifiers will depend on the chosen combination option. For example, classifier $V_{i,j}$ can be defined as follows: $V_{i,j}$ classifies pixel r of the RGB image with the intensity vector $\mathbf{x} = (x_1, x_2, x_3)^T$ as a pixel of category k if and only if

$$\ln f_{k,i}(\mathbf{x}) - \ln f_{s,i}(\mathbf{x}) + \ln f_{k,j}(\mathbf{x}) - \ln f_{s,j}(\mathbf{x}) > 0, \quad (14)$$

for all $s \neq k$.

It is easy to check that the classification condition (14) can be re-formulated as follows: $V_{i,j}$ classifies pixel r of the RGB image with the intensity vector $\mathbf{x} = (x_1, x_2, x_3)^T$ as a pixel of category k if and only if

$$\max_{s=1,7}\left(f_{s,i}(\mathbf{x}) f_{s,j}(\mathbf{x})\right) = f_{k,i}(\mathbf{x}) f_{k,j}(\mathbf{x}). \quad (15)$$

The described combination procedure of classifiers can be generalized and referred to the number of distributions $n > 2$. Thus, for example, we can create the classifier $V_{i,j,l}$

defined as follows: $V_{i,j,l}$ classifies pixel r of the RGB image with the intensity vector **x** as a pixel of category k if and only if

$$\ln f_{k,i}(\mathbf{x}) + \ln f_{k,j}(\mathbf{x}) + \ln f_{k,l}(\mathbf{x}) =$$
$$= \max_{s=1,7}\left(\ln\left(f_{s,i}(\mathbf{x})\right) + \ln\left(f_{s,j}(\mathbf{x})\right) + \ln\left(f_{s,l}(\mathbf{x})\right)\right). \quad (16)$$

In an analogous way, classifier $V_{i,j,l}'$ can be defined for pixel of the TI image with the intensity vector **y**.

But if we want to exploit information from both images and use combination of all 3 distributions, it is possible to define the classifier $W_{1,2,3}$ which classifies pixel r of the RGB image with the intensity vector **x** as a pixel of category k if and only if

$$\ln f_{k,1}(\mathbf{x}) + \ln f_{k,2}(\mathbf{x}) + \ln f_{k,3}(\mathbf{x}) +$$
$$+ \ln g_{k,1}(\mathbf{y}^a) + \ln g_{k,2}(\mathbf{y}^a) + \ln g_{k,3}(\mathbf{y}^a) =$$
$$= \max_{s=1,7}\begin{pmatrix} \ln\left(f_{s,1}(\mathbf{x})\right) + \ln\left(f_{s,2}(\mathbf{x})\right) + \ln\left(f_{s,3}(\mathbf{x})\right) + \\ + \ln\left(g_{s,1}(\mathbf{y}^a)\right) + \ln\left(g_{s,2}(\mathbf{y}^a)\right) + \ln\left(g_{s,3}(\mathbf{y}^a)\right) \end{pmatrix}. \quad (17)$$

Results of experiments performed using classifiers $V_{i,j,l}$, $V_{i,j,l}'$ and $W_{1,2,3}$ provided the classification accuracy estimates presented below.

VII. RESULTS AND CONCLUSIONS

Table I illustrates estimated accuracy of separate classifiers K_i and K_i'. It can be noticed that the best results are obtained using classifiers based on Gaussian distribution. It suggests that the Gaussian distribution is the most adequate model to the real data, of course, remaining within the informal sense of this concept.

TABLE I. OVERALL ACCURACY (%) OF CLASSIFIERS BASED ON SEPARATE IMAGES AND SINGLE PROBABILITY DISTRIBUTION.

Distribution \ Image	RGB	TI
Gaussian	93.03	54.88
Dirichlet	90.65	44.16
Gamma	71.64	32.88

Table II presents estimated accuracy of classifiers $U_{i,j}$ using the combination option 1 that proved to be the most fruitful. It can be noticed that assuming Gaussian distribution for both images is the best variant but usage of Dirichlet distribution for the RGB image data provides considerable accuracy as well.

TABLE II. OVERALL ACCURACY (%) OF CLASSIFIERS USING DATA FUSION AND DIFFERENT PROBABILITY DISTRIBUTIONS.

Classifer $U_{i,j}$	$j=1$	$j=2$	$j=3$
$i=1$	94.96	92.36	91.74
$i=2$	94.41	90.70	91.66
$i=3$	87.28	76.99	66.81

Table III shows overall accuracy estimates obtained using combinations of different distributions for individual images together with data fusion.

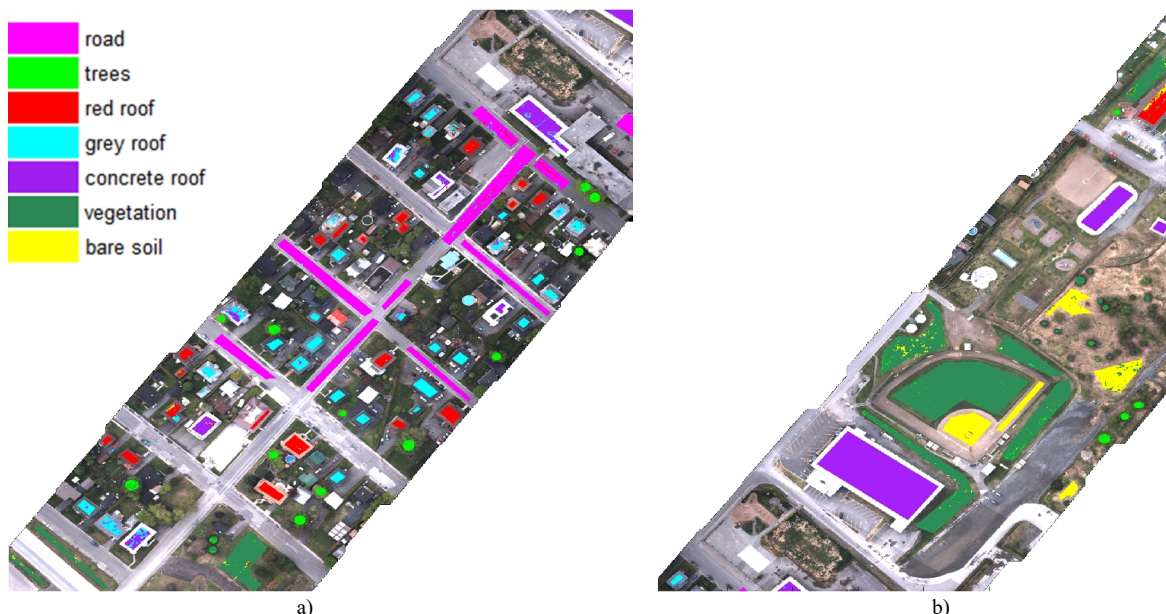a)   b)

Fig. 1. Classification results (classified pixels coloured over the RGB image) of the individual ground truth pixels in the "subset" image using data from both images and combination of all 3 considered distributions: (a) lower part of the "subset" image; (b) upper part of the "subset" image.

TABLE III. OVERALL ACCURACY (%) OF CLASSIFIERS USING DATA FUSION AND COMBINATIONS OF DISTRIBUTIONS.

Distributions\images	RGB	TI	RGB+TI
i=1, j=2	93.29	54.30	95.23
i=1, j=3	93.09	51.88	94.14
i=2, j=3	93.06	52.12	94.56
i=1, j=2, l=3	94.23	54.02	95.79

TABLE IV. CONFUSION MATRIX (%) OF THE CLASSIFIER $W_{1,2,3}$ FOR THE RGB IMAGE USING COMBINATION OF DISTRIBUTIONS.

		Classified category						
		road	trees	red r.	grey r.	concr.	veget.	b.soil
Actual category	road	**96,31**	0,00	1,07	1,69	0,13	0,01	0,78
	trees	0,00	**80,82**	0,00	0,28	0,00	18,91	0,00
	red roof	0,34	0,00	**93,41**	0,64	0,07	0,00	5,55
	grey roof	8,09	0,00	2,08	**84,25**	5,55	0,01	0,02
	concrete r.	0,48	0,00	0,05	2,17	**97,25**	0,00	0,05
	vegetation	0,00	1,99	0,28	0,10	0,00	**94,30**	3,33
	bare soil	0,00	0,08	0,88	0,10	0,34	1,59	**97,01**

TABLE V. CONFUSION MATRIX (%) OF THE CLASSIFIER $W_{1,2,3}$ FOR THE TI IMAGE USING COMBINATION OF DISTRIBUTIONS.

		Classified category						
		road	trees	red r.	grey r.	concr.	veget.	b.soil
Actual category	road	**96,28**	0,04	0,00	2,87	0,53	0,00	0,28
	trees	0,61	**32,35**	0,00	1,20	1,46	54,51	9,87
	red roof	13,86	4,66	**4,75**	70,19	1,19	4,47	0,89
	grey roof	15,05	1,67	0,28	**80,85**	1,75	0,19	0,21
	concrete r.	4,30	31,85	0,00	8,59	**2,99**	41,09	11,18
	vegetation	0,05	22,72	0,00	0,21	0,30	**65,22**	11,49
	bare soil	0,00	11,19	0,00	0,00	0,00	36,32	**52,49**

TABLE VI. CONFUSION MATRIX (%) OF THE CLASSIFIER $W_{1,2,3}$ USING DATA FUSION AND COMBINATION OF DISTRIBUTIONS.

		Classified category						
		road	trees	red r.	grey r.	concr.	veget.	b.soil
Actual category	road	**97,45**	0,00	1,26	1,18	0,10	0,00	0,00
	trees	0,00	**78,94**	0,00	0,03	0,00	21,03	0,00
	red roof	0,23	0,00	**91,94**	2,03	0,03	0,30	5,47
	grey roof	3,25	0,00	1,18	**91,18**	4,40	0,00	0,00
	concrete r.	0,40	0,00	0,07	1,94	**97,38**	0,00	0,20
	vegetation	0,00	1,38	0,16	0,08	0,00	**97,08**	1,29
	bare soil	0,00	0,05	0,00	0,00	0,23	2,87	**96,85**

It is noticed that combination of different distributions for modelling the image data improves classification. The best results are provided by the classifier $W_{1,2,3}$ using data from both images and combination of all 3 considered distributions (see Fig. 1, corresponding confusion matrices are presented in Table IV–Table VI).

Such result is somewhat problematic to interpret. Further study might be necessary to provide proper explanation. The first hypothesis is that combination of different distributions might compensate for the imperfections of models based on individual distributions.

REFERENCES

[1] A. Lorencs, I. Mednieks, J. Sinica-Sinavskis, "Classification of multisensor images with different spatial resolution", *Elektronika ir Elektrotechnika*, vol. 21, no. 5, pp. 81–85, 2015. [Online]. Available: http://dx.doi.org/10.5755/j01.eee.21.5.13333
[2] S. Kotz, N. Balakrishnan, N. L. Johnson, *Continuous Multivariate Distributions, Vol. 1: Models and Applications*. Wiley, 2000.
[3] N. Solaro, "Random variate generation from Multivariate Exponential Power distribution", *Statistica & Applicazioni*, vol. 2, no. 2, 2004.
[4] 2014 IEEE GRSS Data Fusion Contest. [Online]. Available: http://www.grss-ieee.org/community/technical-committees/data-fusion/
[5] B. Storvik, G. Storvik, R. Fjortoft, "On the combination of multisensor data using meta-Gaussian distributions", *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2372–2379, 2009. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2009.2012699
[6] A. Voisin, V. A. Krylov, G. Moser, S. B. Serpico, J. Zerubia, "Supervised classification of multisensor and multiresolution remote sensing images with a hierarchical copula-based approach", *IEEE T. Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3346–3358, 2014. [Online]. Available: http://dx.doi.org/10.1109/TGRS.2013.2272581