CLASSIFYING MICROWAVE RADAR IMAGES USING DECISION BASED DATA FUSION

Markus Törmä[a], Marcus Engdahl[b]
[a]Institute of Photogrammetry and Remote Sensing, Helsinki University of Technology, Otakaari 1, Espoo, Finland - Markus.Torma@hut.fi
[b]Directorate of Earth Observation Programmes, ESA-ESRIN, Via Galileo Galilei, Casella Postale 64, 00044 Frascati, Italy - Marcus.Engdahl@esa.int

KEY WORDS: Classification, Land Cover, Fusion, SAR

ABSTRACT:

Four different data fusion methods for classifying SAR images with different spatial resolutions were tested and compared. Test area was Helsinki region in Southern Finland and test data 14 ERS-1/2 Tandem pairs. The best method was a method where a posteriori probabilities of lower spatial resolution classification are used as a priori probabilities of higher resolution classification. The increase of overall accuracy was 7-14 %-units depending on date and 10.4% on average when compared to original Tandem pairs. Median filtering increased classification accuracy, but not that much when data fusion methods were used. This means that the need of spatial filtering can be at least partially compensated using data fusion of different spatial resolution images.

1. INTRODUCTION

Many governmental institutions have a continuing requirement to form and implement laws and policies that involve existing or future land cover and land use. Remote sensing offers means to provide information about the environment. For example, the purpose of CORINE Programme is to gather information relating to the environment for the European Union. In order to determine and assess the effects of Community's environment policy, it is needed to have a proper understanding concerning the different features of the environment like the state and geographical distribution of individual environments and natural areas, the quality and abundance of water resources, land cover and soil state, and natural hazards (Heymann et.al., 1994).

CORINE land cover classification is produced using optical satellite images like Landsat ETM. Unfortunately weather conditions limit the use of optical data. For example, here in Finland summer is usually quite cloudy, there are usually only few days when large area of Finland is cloud-free, and during winter there is dark also daytime. Finnish IMAGE2000 satellite image mosaic consisted of 36 Landsat ETM images. The target year was 2000 but only 12 images were taken year 2000 due to cloudiness (Härmä et.al., 2004).

Due to the relative insensitivity to weather, SAR images are an interesting alternative to produce information about land cover and land use. One commercially important application is cellular network planning, where topographic (DEM) and morphographic data are needed. SAR interferometry is used to produce the topographic information. SAR intensity images and coherence image produced in interferometric process are used to provide information about the land cover. Morphographic classes are defined as different land cover types, according to how they interact and attenuate electromagnetic radiation. The most common morphographic categories are urban, suburban, rural, water, open, and forest (Hyyppä et.al., 1999). So, the aim here is to acquire required spatial information using only two SAR images taken in slightly different places.

Traditional and difficult problem when classifying SAR images is speckle, in other words image contains random noise due to the measurement process. Due to speckle, the statistics of different land cover classes are rather similar, making the statistical classification difficult. Speckle can be decreased by different kind of filters, but developed methods can be difficult to use due to many parameters that can be unknown. Simple alternative is performing average filtering. As the area of averaging increases class histograms start to resemble normal distribution that is good from statistical pattern recognition point of view, but spatial resolution decreases at the same time.

This problem is studied from decision based data fusion point of view in the article. The main principle is that by merging several different classification results made using images with different spatial resolutions, it is possible to acquire better classification than by using any individual classifications. First, ERS SAR-images are averaged using different sizes of filter and then these images are classified. Then different methods are tested to merge these different classifications to form a better one.

The aims of this research are twofold. First, several rather simple decision based data fusion methods are introduced and their performance compared. Second, due to large number of SAR images, 14 ERS-1/2 Tandem pairs, the effect of environmental conditions to classification results can be assessed.

2. DECISION MAKING AND DATA FUSION

Interpretation of remote sensing data can be divided to two approaches, modelling and classification. Modelling means that some geophysical parameter is estimated from remote sensing data, like soil moisture or forest stem volume. The aim of classification is to divide measurements to discrete groups or classes according to their similarities. This requires that for each measurement we make decision about the most proper or likely class.

2.1 Decision making

A common means to perform classification is to use statistical pattern recognition framework. There are several different approaches to classification but the Bayes rule is commonly

utilized. Bayes rule measures the a posteriori probability $P(\omega_j|x)$ which feature vector x belongs to class j as (Devivjer and Kittler, 1982):

$$P(\omega_j \mid x) = \frac{p(x \mid \omega_j)P_j}{p(x)}, \quad (1)$$

where P_j is the a priori probability of class j, $p(x|\omega_j)$ is the value of density function of class j and $p(x)$ is the mixture density function of x defined as

$$p(x) = \sum_{j=1}^{c} p(x \mid \omega_j)P_j, \quad (2)$$

where c is number of classes.

Feature vector x can be classified to class j if the a posteriori probability of that class is larger than the a posteriori probabilities of other classes. In that case the decision rule is called Bayes rule for minimum error. If it is thought that uncertain classification should not be done, rejection threshold λ_r can be used. In that case decision rule becomes

$$\begin{aligned} &\omega(x) = \omega_i \quad if \\ &P(\omega_i \mid x) = \max_{j=1}^{c} P(\omega_j \mid x) \geq 1 - \lambda_r \\ &\omega(x) = \omega_0 \quad if \\ &\max_{j=1}^{c} P(\omega_j \mid x) < 1 - \lambda_r \end{aligned} \quad (3)$$

In other words, classification decision is not made if the largest a posteriori probability is less than 1 - rejection threshold λ_r. Cost function $\lambda(\omega_i|\omega_j)$ can be used to make some classification decisions more important than the others. The purpose of cost function is to punish wrong decisions so it can be thought to weight different decisions.

The density function measures the distance between feature vector x and class j. Remote sensing data is often normally distributed, especially optical data, so normally distributed density function

$$p(x \mid \omega_j) = (2\pi)^{-d/2} \left|\Sigma_j\right|^{-\frac{1}{2}} e^{\left[-\frac{1}{2}(x-\mu_j)^T \Sigma_j^{-1}(x-\mu_j)\right]} \quad (4)$$

can be used. Σ_j is the covariance matrix of class j, μ_j the mean vector of class j and d is the dimensionality of feature space.

2.2 Alternatives for data fusion

Data fusion can be performed on different levels (Pohl and van Genderen, 1998):

- Pixel based fusion means that the measurements or measured physical parameters have been fused. In other words, the feature vector is combined directly from different data sources.
- Feature based fusion means that features have been extracted from different data sources using e.g. image segmentation. In this case the features can be e.g. size, shape and average intensity level of areas. These features form feature vectors describing the extracted objects.
- Decision based fusion means that the objects have been identified from individual data sources and then these interpretation results are combined using e.g. rules to reinforce common interpretation.

2.3 Decision based data fusion methods

One way to perform decision based data fusion is to use the class labels of individual classifications and making some kind of majority decision. This majority decision can be due to simple majority voting (Ho et.al., 1994) or consensus builder (Liu et.al., 2002). Other examples of this approach include methods for making the decision by ranking the class labels of individual classifications (Ho et.al., 1994), using special neural network classifier to classify samples if their statistical and neural network classifications disagree (Kanellopuolos et.al. 1993), or using the classification of expert system as input to neural network classifier (Liu et.al., 2002). This approach could be called hard decision based data fusion.

Soft ecision based data fusion is another alternative. In that case probabilities or other kind of measures which pixel belongs to certain class are used. This latter approach is used in this study. All implemented data fusion methods use the output of Bayes decision rule, i.e. a posteriori probabilities, as their input.

2.3.1 Maximum a posteriori probability: In this case classification decision is based on the maximum a posteriori probability of different classifications cl. Decision rule is

$$\begin{aligned} &\omega(x) = \omega_i \quad if \\ &P(\omega_i \mid x) = \max_{k=1}^{cl} \max_{j=1}^{c} P(\omega_j \mid x_k). \end{aligned} \quad (5)$$

This rule corresponds to using fuzzy union operator to combine different fuzzy sets (Zimmermann, 2001). The drawback is that information about the reliability of different classifications is not used.

2.3.2 Maximum joint a posteriori probability: A joint a posteriori probability of classification decision can be computed as (Swain, 1978)

$$P(\omega_i \mid x) = \prod_{k=1}^{cl} P(\omega_i \mid x_k), \quad (6)$$

where $P(\omega_i|x)$ is the a posteriori probability of class j computed using feature vector x_k from data source k. In this case it is assumed that the individual classifications are independent from each other, but this is not always the case in real life. Another drawback is that information about the reliability of different classifications is not used. This can be incorporated by using accuracy coefficient α_k so that the equation of joint a posteriori probability is

Nr	Date	Track/Frame	Base (m)	Air temperature C° (9:00 UTC)	Wind speed (direction) (m/s, deg.)	Snow depth (cm)	Precipitation between images (mm)
1	17/18 July 1995	408/2385	-2	19.5 17.1	8 (140) 5 (190)	0 0	3.4
2	21/22 August 1995	408/2385	-72	17.1 21.1	3 (330) 5 (240)	0 0	0
3	9/10 September 1995	179/2385	-28	11.5 10.3	6 (30) 3 (30)	0 0	14.9
4	25/26 September 1995	408/2385	239	13.5 11.7	8 (190) 10 (210)	0 0	1.2
5	14/15 October 1995	179/2385	-221	8.6 6.2	6 (290) 1 (90)	0 0	0
6	30/31 October 1995	408/2385	-49	5.2 -1.7	2 (260) 5 (340)	0 0	0.1
7	8/9 January 1996	408/2385	-29	-6.5 -5.0	2 (120) 3 (170)	13 13	0
8	12/13 February 1996	408/2385	85	-14.9 -10.8	5 (100) 1 (120)	16 19	1.4
9	2/3 March 1996	179/2385	-76	-4.2 -5.5	4 (10) 3 (50)	32 32	0.1
10	18/19 March 1996	408/2385	80	-3.2 -4.4	3 (350) 1 (110)	39 38	0
11	6/7 April 1996	179/2385	37	5.5 7.6	4 (220) 2 (130)	38 (wet snow) 32 (wet snow)	0
12	22/23 April 1996	408/2385	-58	14.7 10.0	5 (240) 3 (50)	0 0	0
13	15/16 June 1996	179/2385	48	15.6 14.9	7 (330) 6 (340)	0 0	0
14	20/21 July 1996	179/2385	188	16.0 14.6	4 (70) 5 (40)	0 0	0.1

Table 1. The environmental conditions of acquired ERS-1/2 SAR images. Images have been taken approximately 10:35 UTC (Pulliainen et.al., 2003).

$$P(\omega_i \mid x) = \prod_{k=1}^{cl} P(\omega_i \mid x_k)^{\alpha_k}, \qquad (7)$$

Accuracy coefficient α_k can be related to classification accuracy (e.g. the estimated classification accuracy of x) or spatial resolution giving lower coefficient to classification that is based on lower resolution images.

2.3.3 A priori probabilities from lower quality probabilities or other data: A priori probability of Bayes rule represents our knowledge about the truth of the hypothesis before we have analysed the current data. This knowledge could be acquired using old classification or derived from statistical data. One way to estimate a priori probability is to compute the a posteriori probabilities of lower resolution classification and use these as a priori probabilities in higher resolution classification (Schneider et.al., 2003, Törmä et.al., 2004). In the case of remote sensing, this lower resolution could mean that the spatial, radiometric or spectral resolution is lower or otherwise less suitable for classification problem at hand.

2.3.4 Dempster-Shafer theory of evidence: The mathematical theory of evidence is a method to combine different data sources to provide a joint inference concerning the correct classification. The idea is to assign a so-called mass of evidence to various labeling propositions for a feature vector. Each vector has its own mass of evidence describing the likelihood of different classes as well as some indication about the uncertainty about the labeling. For each labeling proposition, values of support and plausibility are computed. Support is considered to be the minimum amount of evidence in favor of a particular labeling for a pixel whereas plausibility is the maximum possible evidence in favor of the labeling. The difference between the measures of plausibility and support is called the evidential interval, the true likelihood that the label under consideration is correct for the pixel is assumed to lie somewhere in that interval. Orthogonal sum is used to combine evidence from different sources. After the orthogonal sum has been applied, the user can then compute the support for and the plausibility of each class for a feature vector. The final decision is based on the support and plausibility. In the simplest case a maximum support rule can be used (Richards, 1993).

3. SAR IMAGES AND GROUND TRUTH

Study site is the Helsinki region in Southern Finland, covering some 50 km x 50 km area. ERS-1/2 images were acquired during summer 1995 - summer 1996 consists of 14 Tandem image pairs. The temporal separation between image acquisitions is 24 hours in a Tandem pair. Measurement parameters of ERS are C-band, frequency 5.3 GHz, polarization VV, incidence angle 23 degrees and spatial resolution about 30 m (Kramer, 1996). Two 5-look backscattered intensity images and an interferometric coherence image estimated with a 5x5 pixel Gaussian window were produced from each Tandem pair. All image data was orthorectified to national coordinate system using an INSAR DEM. Weather conditions within image acquisition times were also recorded. Table 1 presents the dates and environmental conditions of image acquisitions. The intensity and coherence images were scaled to 8-bit integers.

Chosen land cover and use classes are presented in table 2 with number of ground truth samples per class for different spatial resolutions. Training areas for classes were determined using Vantaa city regional map made by Vantaa city authorities and National Land Use and Forest Classification made by National Land Survey. Regional map and LUFC were combined to one map with 20 m pixel. In order to decrease the effect of georeferencing errors, 2 pixels were removed from borders of map areas. Then the coordinates of pixels were written to ascii-files and systematically sampled so that the training data of each class would contain about 1000 pixels at 20 m spatial resolution.

Class	80m	40m	20m
1. Water	74	287	1119
2. Agricultural and other open area	39	252	1000
3. Dense forest, stem volume over 100 m3/ha	61	265	1030
4. Sparse forest, stem volume 50-100 m3/ha	58	261	1000
5. Single story houses	48	191	779
6. Multi-story houses	40	164	705
7. Industrial area	62	239	1008
SUM	382	1659	6641

Table 2. Chosen land cover and use classes with number of ground truth samples per class for different spatial resolutions.

4. EXPERIMENTS

Statistical classifications were made using Bayes decision rule with Maximum Likelihood density function estimation method. Classes were supposed to be normally distributed. Classification errors were estimated using resubstitution and holdout methods. In resubstitution method the same set is used as training and test set and in holdout method data is randomly divided to training and test sets (Devivjer and Kittler, 1982). In order to decrease the effect of random division, all classifications were repeated three times and the mean values of accuracy measures computed. Because resubtitution method is positively biased and holdout negatively biased, final error estimate was mean error of these two estimates. Error matrix was used to compare the classification results and reference data. Several accuracy measures like Overall accuracy, Producer's accuracies of individual classes, and User's accuracies of individual classes were computed from error matrix (Lillesand and Kiefer, 1994).

Data fusion methods presented in chapter 2.3 were implemented and used the output of the statistical classifier, i.e. a posteriori probabilities, as their input:

- Maximum a posteriori probability (DF1) method was implemented as presented in chapter 2.3.1.
- Maximum joint a posteriori probability (DF2) method has four different versions according were and how the reliability measures were used. In case A, each individual classification has same influence to final decision. In cases B, C and D, the individual classifications are weighted using the probability of correct classification or fixed values. In case B, the density function value of class is weighted by the estimate of the probability of correct classification that is same as the a posteriori probability of that class. In case C, the density function value of class is weighted by the maximum a posteriori probability of decision. In case D, fixed values were used and they were 0.5 for 80 m spatial resolution images, 0.7 for 40 m and 0.9 for 20 m.
- A priori probabilities from lower quality probabilities or other data (DF3) method used the a posteriori probabilities of lower spatial resolution classification as a priori probabilities of higher spatial resolution classification.
- Dempster-Shafer theory of evidence (DF4) method was implemented according to Richards (1994). There were two versions differing how the uncertainty of labeling was defined. Case A used a posteriori probabilities of individual classifications, as case B had fixed values 0.5 for 80 m spatial resolution images, 0.7 for 40 m and 0.9 for 20 m.

5. RESULTS

5.1 Effect of spatial resolution to statistical classification

Figure 1 illustrates the effect of spatial resolution to the classification accuracy. Horizontal axis corresponds to ERS-1/2 Tandem pair (table 2) and vertical axis the overall accuracy. Solid red line represents the accuracies acquired using 20 m spatial resolution images, solid blue line with "x" 40 m images, solid green line with "o" 80 m images, dashed red line 20 m median filtered images, dashed blue line with "x" 40 m median filtered images, and dashed green line with "o" 80 m median filtered images.

From classification accuracy point of view, averaging of SAR images increases overall accuracies. This is due that the random component of class statistics is suppressed. The drawback is that the spatial resolution decreases at the same time. When the overall accuracies of 20 and 40 m spatial resolution images are compared, it is noticed that the difference in accuracies is 0-6 %-units depending on Tandem pair and 2.9 %-units on average. It seems that differences are largest during summer or winter in rather dry conditions. In other words, moisture decreases the differences of classification accuracies obtained using different spatial resolutions. When 20 and 80 m, or 40 and 80 m images are compared, the differences are larger. They are on average 7.5 %-units between 20 and 80 m images and 4.6 %-units between 40 and 80 m images. In these cases there are no clear correlation between difference and image acquisition times.

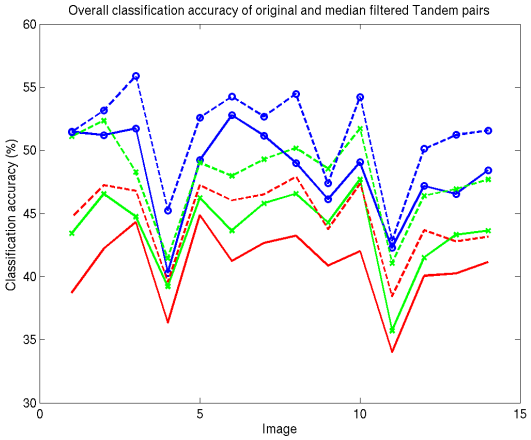

Figure 1. The effect of spatial resolution to the classification accuracy

The classification accuracies depend quite heavily on acquisition time and their environmental and weather conditions. The worst classification accuracies were acquired using image 4 in relatively windy (wind direction about same in different dates) conditions and image 11 in wet snow conditions. The dates of best classification accuracies varied more and depended on spatial resolution. The best overall accuracy, 44.9%, for 20 m images was obtained using image 5, 46.6% for 40 m spatial resolution using image 8, and 52.8% for 80 m spatial resolution using image 6.

The median filtering of images increases classification accuracy. The increase is 2-6 %-units (average 3.8 %-units) in the case of 20 m spatial resolution images, 2-7 %-units (average 4.3 %-units) in the case of 40 m images and 0-5 %-units (average 2.9 %-units) in the case of 80 m images. The best overall accuracy, 47.9%, for 20 m images was obtained using image 8, 52.3% for 40 m spatial resolution using image 2 and 55.9% for 80 m spatial resolution using image 3.

5.2 Comparison of data fusion methods

Figure 2 illustrates the classification accuracies of different data fusion methods as function of time. Horizontal axis corresponds to ERS-1/2 Tandem pair and vertical axis overall accuracy. Solid red line represents the accuracies of statistical classifier and 20 m spatial resolution images, solid green line with "x" data fusion DF1, solid blue line with "o" DF2C, dashed green line with "x" DF3 and dashed blue line with "o" DF4A.

Data fusion increased classification accuracies in each case. The best method was DF3, the increase of overall accuracy is 7-14 %-units depending on date and 10.4% on average. DF2C was the second best; the increase is 4-10 %-units and 7.2 %-units on

average. The increase of accuracy was usually smaller during spring or autumn. The results of different versions of DF2 were close to each other, but DF2C was usually the best. The differences between versions of DF4 were larger. The worst dates were same as before. The best dates varied between different methods. The best overall accuracy, 55.0%, was obtained using DF3 and image 2.

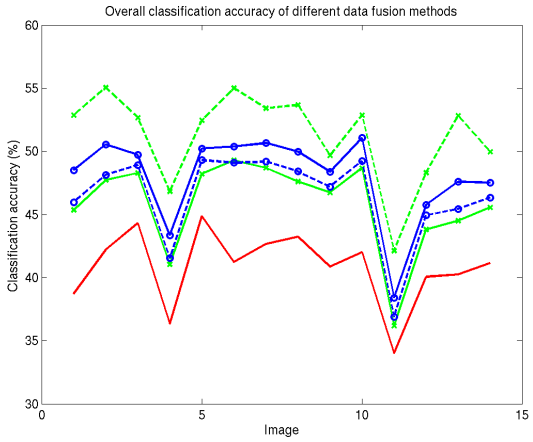

Figure 2. The classification accuracies of different data fusion methods as function of time

Median filtering increased classification accuracy somewhat. The increase was larger, varying 2-6 %-units, for DF1 and DF4A and smallest, varying 0-3 % units, for DF3. The best overall accuracy, 56.5%, was obtained using DF3 and image 2. This means that the need of spatial filtering can be at least partially compensated using data fusion of different spatial resolution images.

5.3 Effect of season

The seasonal comparison was made by computing seasonally averaged Tandem pairs. Images 1, 2, 13 and 14 were temporally averaged to form summer Tandem pair, images 3, 4, 5, and 6 autumn Tandem pair, images 7 and 8 winter Tandem pair and images 9, 10, 11 and 12 spring Tandem pair. Also, all 14 Tandem pairs were temporally averaged to form one Tandem pair. Figure 3 illustrates the classification accuracies of different data fusion methods as function of season. Horizontal axis corresponds to seasonal Tandem pair (1: whole year, 2: autumn, 3: winter, 4: spring and 5: summer) and vertical axis overall accuracy. Solid red line represents the accuracies of statistical classifier and 20 m spatial resolution images, solid green line with "x" data fusion method DF1, solid blue line with "o" DF2C, dashed green line with "x" DF3 and dashed blue line with "o" DF4A.

Temporal averaging increased classification accuracy but surprisingly little. The overall accuracies were about the same for best seasonal Tandem pairs and median filtered Tandem pairs for DF1, DF2C and DF4A. The increase was larger in the case of DF3. When temporal averaging included all Tandem pairs, accuracies were much better. The best overall accuracy, 66.1%, was obtained using DF2C. In the case of DF3, the increase of classification accuracy was smaller and actually the overall accuracy of summer Tandem pair was higher (60.4%) than the Tandem pair of whole year (58.3%).

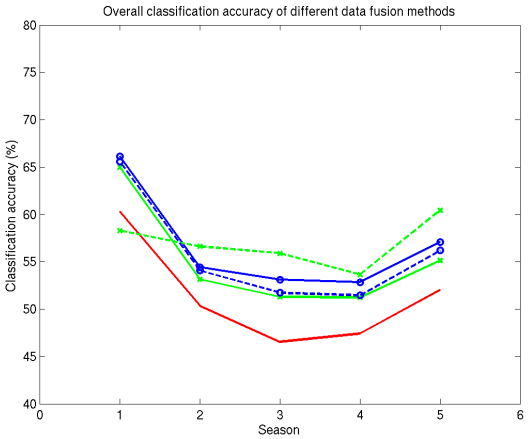

Figure 3. The classification accuracies of different data fusion methods as function of season.

5.4 Individual classes

The classwise accuracies varied a lot. Table 3 presents the best producer's and user's accuracies for each class, used data fusion method and corresponding image. Water and agricultural and other open areas were classified rather well, dense forest and industrial moderately and other classes rather poorly. The best data fusion method seems to be DF3. There was some correlation between season and class accuracies. Producer's accuracies of class Water were smaller during snow-season than other seasons. In the case of Agricultural and other open areas, producer's accuracies were larger during dry snow. The user's accuracies of classes Water and Dense forest were highest during autumn. The user's accuracies of classes Agricultural and other open areas and Single story houses were largest during winter and Multi-story house the best accuracy was achieved just after snowmelt.

Class	Producer's accuracies	User's accuracies
1. Water	98%, DF3, 5,14	95%, DF1, 6,14
2. Agricultural and other open area	78%, DF3, 10	93%, DF1, 9
3. Dense forest, stem volume over 100 m3/ha	61%, DF3, 8	73%, DF3, 3
4. Sparse forest, stem volume 50-100 m3/ha	44%, DF3, 3	54%, DF3, 8
5. Single story houses	40%, DF2A, 1	33%, DF3, 9
6. Multi-story houses	36%, DF3, 13	56%, DF4A, 12
7. Industrial area	65%, DF3, 6	57%, DF3, 13

Table 3. The best producer's and user's accuracies for each class, used data fusion method and corresponding image.

Median filtering increased classwise accuracies somewhat, the increase varies 1-3 % units depending on accuracy measure and date. The most notable difference was class Single story house, in which case user's accuracy increases 12 %-units. The advantage of medium filtering is that variation of classification accuracy decreases between worst and best accuracies.

5.5 Mixing of classes

The mixing of classes to other classes was studied using error matrices. Usually mixing happened following way:

- Water: mixed with Dense forest and Sparse forest
- Agricultural and other open areas: Multi-story houses and Sparse forest
- Dense forest: Sparse forest and Water
- Sparse forest: Dense forest and Single story houses
- Single story houses: Multi-story houses and Sparse forest
- Multi-story houses: Industrial and Single story houses
- Industrial: Multi-story houses and Single story houses

As expected, forest classes were mainly mixed with each other and build-up classes with each other. Main exception was that single story houses and sparse forest was mixed with each other. Water was mainly mixed with forest classes. Greatest surprise was that agricultural and other open areas were quite heavily mixed with multi-story houses. There were little differences between different data fusion methods. Largest differences were between statistical classification of 20 m spatial resolution images and data fusion methods. In statistical classification, agricultural and other open areas were more mixed with build-up areas.

5.6 Merging of classes

The effect of merging of classes from seven classes to four classes was tested using image 6. Four classes were water, agricultural and other open areas, forest (classes dense and sparse forest), and build-up area (classes single story houses, multi-story houses and industrial). Overall classification accuracies increased quite a lot, as expected. This increase was about 20 %-units, from 18.8 %-units in case of statistical classification of 20 m spatial resolution images to 23.7 %-units in case of DF2B. The best classification accuracy, 77.7%, was achieved with DF3. Mixing of classes was mainly as water was mixed with forest, agricultural with build-up, forest with build-up, and build-up with forest. There were no differences with different data fusion methods in this sense.

6. CONCLUSIONS

Four different data fusion methods for classifying SAR images with different spatial resolutions were tested and compared. The best method was DF3 where a posteriori probabilities of lower spatial resolution classification are used as a priori probabilities of higher resolution classification. The increase of overall accuracy was 7-14 %-units depending on date and 10.4% on average when compared to original Tandem pairs. The increase of accuracy was usually smaller during spring or autumn. Median filtering increased classification accuracy, but not that much when data fusion methods were used. This means that the need of spatial filtering can be at least partially compensated using data fusion of different spatial resolution images.

ACKNOWLEDGMENT

The authors wish to thank ESA for providing the SAR data through ESA Announcement of Opportunity studies AOT-SF.301 and A03-277 as well as the town of Vantaa for the high-resolution aerial color orthophotos.

REFERENCES

Devivjer P., Kittler J., 1982. *Pattern Recognition - A Statistical Approach.* Prentice-Hall, 1982.

Ho, T., Hull, J., Srihari, S., 1994. Decision combination in Multiple Classifier Systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 1, January 1994, pp. 66-75.

Hyyppä, J., Törmä, M., Engdahl, M., 1999. Deriving morphographic classification for cellular network planning using SAR interferometry. *IGARSS'99*, Hampuri, 28.6.-2.7.1999. 1999, IEEE, s. 2176-2178.

Härmä, P., Teiniranta, R., Törmä, M., Repo, R., Järvenpää, E., Kallio, M., 2004. The Production of Finnish Corine Land Cover 2000 Classification. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey, Vol. XXXV, part B4, pp. 1330-1335.

Jeon, B., Landgrebe, D., 1999. Decision fusion approach for multitemporal classification. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, May 1999, pp. 1227-1233.

Kanellopuolos, I., Wilkinson, G., Megier, J., 1993. Integration of neural network and statistical image classification for land cover mapping. *Proceedings of the International Geoscience and Remote Sensing Symposium* IGARSS'93, Kogakuin University, Tokyo, Japan, 18-21 August. s. 511-513.

Kramer, H., 1996. *Observation of the Earth and Its Applications*. 3rd enlarged edition, Springer.

Lillesand T., Kiefer R., 1994. *Remote Sensing and Image Interpretation.* John Wiley & Sons, Inc, 1994.

Liu, X-H., Skidmore, A., Van Oosten, H., 2002. Integration of classification methods for improvement of land-cover map accuracy. *ISPRS Journal of Photogrammetry and Remote Sensing*, 56, pp. 257-268.

Pohl, C., van Genderen, J., 1998. Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications. *International Journal of Remote Sensing*, 19(5), pp. 823-854.

Pulliainen, J., Engdahl, M., Halikainen, M., 2003. Feasibility of multi-temporal interferometric SAR data for stand-level estimation of boreal forest stem volume. *Remote Sensing of Environment*, vol. 85, pp. 397-409.

Richards, J., 1993. *Remote Sensing Digital Image Analysis*. 2nd ed., Springer, 1993.

Schneider, A., Friedl, M., McIver, D., Woodcock, C., 2003. Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, vol. 69, no. 12, December 2003, pp. 1377-1386.

Swain, P., 1978. Bayesian classification in a time-varying environment. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, no. 12, pp. 879-883.

Törmä, M., Patrikainen, N., Luojus, K., Lumme, J., Pyysalo, U., 2004. Tree Species Classification Using ERS SAR and MODIS NDVI Images. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey, Vol. XXXV, part B7, pp. 927-932.

Zimmermann, H.-J., 2001. *Fuzzy Set Theory and Its Applications.* 4th ed. Kluver Academic, p. 17.