OPTICAL IDENTIFICATION OF CONSTRUCTION AND DEMOLITION WASTE BY USING IMAGE PROCESSING AND MACHINE LEARNING METHODS

Katharina Anding[1], Elske Linß[2], Hannes Träger[2], Matthias Rückwardt[1] and André Göpfert[1]

[1]Faculty of Mechanical Engineering, Department of Quality Assurance and Industrial Image Processing, Ilmenau University of Technology, Ilmenau, Germany
[2]Faculty Civil Engineering, Bauhaus-University of Weimar, Weimar, Germany

Abstract − This paper discusses the possibility of the optical identification of recycled aggregates of construction and demolition waste (CDW) as basis of an innovative sorting method on the field of processing of CDW. The first target was to find suitable identification attributes for the differentiation of aggregates, which are difficult to separate. For the investigations images of the given aggregate classes were captured and analysed by algorithms of image processing and machine learning. The interdependencies between dataset character, feature vector, type of the selected classifier and parameter settings of classifier are very complex and they were analyzed in this paper.

Keywords: optical identification, machine learning, construction and demolition waste (CDW)

1. INTRODUCTION

Construction and demolition waste (CDW) are the biggest waste flow in Germany. There was an amount of 72.1 Mio t CDW in the year 2004 [1]. The recycling rate amounts to 70 % (49.6 Mio. t). Certainly the recycling rate depends on the composition and heterogeneity of material (figure 1).

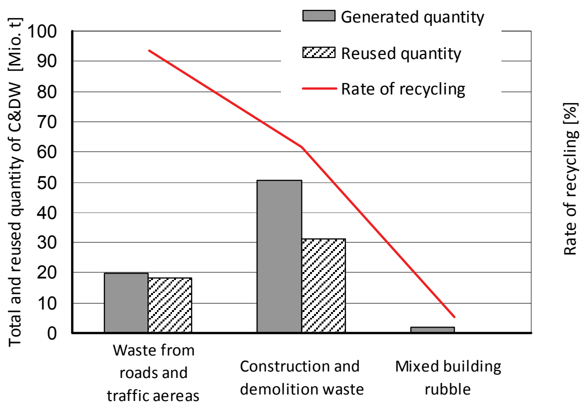

Fig. 1. Reuse of different kinds of CDW

A significant decrease can be observed in the recycling rate with increasing heterogeneity of the recycled material.

For recycled masonry aggregates and recycled mixed aggregates the lowest recycling rates are found because of the high heterogeneity and the mineral admixtures. Therefore the reuse of these materials is very difficult.

Regarding to application of C&D aggregates, most of them are used in road pavements and earth works, not really substituting the natural aggregate applications. Only a very small part of around 5 % flows back in the production of recycling concrete.

The down cycling scenario is also common in other countries [2], [3]. It is not considered sustainable because the use of land will be critical in the future and land filling should be avoided [4].

The recycling industry of building materials is dominated by simple technologies. For instance the single-stage crushing is used with advance sieving and separation of reinforcement steel by over belt magnetic separator. For the processing of building materials sorting processes are only used for the separation of light components until now.

These technologies are not able to separate the incidental mixed aggregates. They are suitable in no way for “new building materials” including connected building materials, which will use more and more in building industry.

CDW from building construction are heterogeneous mixtures of brick, mineral bounded building materials (concrete, sand-lime-brick, aerated concrete, lightweight concrete), mortar, plaster, insulation material, wood and plastic etc. Sorting analyses of crushed CDW confirm the variety in recycling materials. Analyses of the density show a large range of bulk density. The water adsorption is much higher and the grain strength lower than for natural aggregates. The low quality of C&D aggregates and their variability seems to be the most important aspect to limit its application in concrete. In fact, the composition and physical properties of C&D aggregates are variable in a wide range.

The heterogeneity prevents the profitable reuse. Therefore it is necessary to reduce the heterogeneity. But this is possible only by a multi-stage process with several classify and sorting steps.

It is indispensable to separate the CDW mixtures to establish a reliable and demanding reuse. This is the basis for the development of specific products which based on the characteristical properties of the materials. And it is also the basis for the return of pure material as secondary material in the production of primary material.

The aim is the realization of real closed cycles and a high standard of quality in recycling.

State of the art

As in other sectors of recycling, for example the glass or plastic recycling, the sensor-based sorting has been becoming more interesting in the recycling of building materials and sorting of minerals in the last years. There are mainly used methods with optical, magnetic, NIR or X-Ray sensors.

The application of automatic sensor sorting in the areas of mining and recycling is successful in Europe and will increase in the following years. The benefits are the increase of the end product value and the cost reduction of downstream handling steps in the processing [5], [6].

Aim of Investigations

The first investigations are focussed on the optical differentiation of phenotypically similar building materials like concrete, aerated concrete, lightweight concrete and also porous and dense brick. First investigations were done on new, not used building materials, which were crushed. Their parameters bulk-density, porosity and water adsorption are shown in Table 1.

Table 1. Parameters of samples

Sample	Bulk density in Mg/m³	Porosity in %	Water adsorption in %
concrete			
2	2,4605	9,3	3,0
3	2,4437	10,0	4,7
4	2,2593	18,2	6,3
5	2,6717	10,4	3,4
6	2,4676	8,6	3,7
7	2,4082	10,2	4,3
8	2,4110	11,2	6,0
10	2,4808	10,5	5,0
aerated concrete			
1	0,7512	70,5	74,1
2	0,7685	70,0	99,9
3	0,8044	68,9	111,0
4	0,8636	66,6	82,4
5	0,8855	66,0	67,7
6	0,7763	69,8	85,3
7	0,5979	76,7	116,9
lightweight concrete			
1	0,8707	65,4	43,2
2	0,6716	75,2	29,8
3	0,9190	66,3	28,1
4	0,8250	65,8	42,9
5	1,3857	47,8	17,4
6	1,2988	52,0	25,3
7	1,2386	53,3	23,1
8	1,8654	30,1	10,9
9	1,4314	49,7	13,2
porous brick			
3	1,6836	40,2	23,6
5	1,7257	37,2	23,0
dense brick			
4	2,2226	18,0	8,8

In addition an optical solution for determination of building material classes was investigated by using methods of image processing and machine learning. Several optical attributes were found, which have discriminatory power to classify the chosen materials.

Several classification algorithms of supervised machine learning were tested on different feature vectors as a numerical representation of objects of the given dataset. The different feature vectors were built by using feature selection methods, especially filter methods like Information Gain [7] and Chi Squared [8], [9]. As a result the best differentiating features and the most qualified classifiers were attained for solving this optical identification task of building materials.

2. REALIZATION OF INVESTIGATIONS

A precondition for a satisfying technical performance of the automated analysis is a good analyzable image. So images were taken of 35 different material samples (8 concretes, 7 aerated concretes, 9 lightweight concretes and 3 bricks). The samples were captured by a RGB matrix camera (see figure 2). A combination of incident and transmitted light was chosen for capturing colour images. The lighting device consists of three LED-light lines and a light table to visualize the class specific in an optimal way. All images were taken under constant conditions.

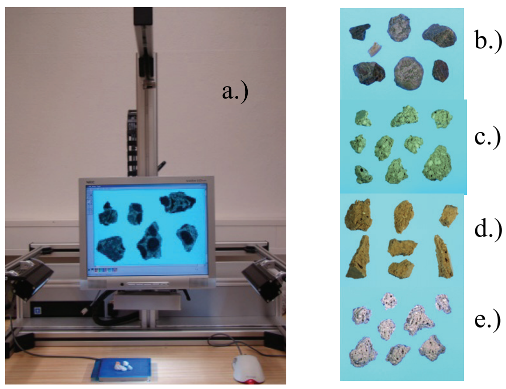

Fig. 2. a.) Image system „QI Inspector“ [10] and images of samples b.) concrete, c.) lightweight concrete, d.) brick, e.) aerated concrete

A blue foil was installed on the light table to separate the particles as good as possible from the background. Almost 1000 images of particles out of each material class were taken to realize a good statistical comparison. It means that more than 100 particles per sample were captured. The particle sizes were in the range of 4 to 8 mm.

2.1. Feature Extraction

After segmentation and transformation from the RGB to HSI colour space a feature vector for every object has been calculated. 32 colour and texture features were used like the mean value per channel and features calculated from the co-occurrence matrix like energy, homogeneity and contrast per each HSI-channel. 92 scale and rotation-invariant shape features were also calculated like modified Fourier descriptors.

The used feature algorithms are part of the machine vision software Halcon and were described in the release notes for MVTec HALCON 8.0.3 [11]. Here three of the 188 used feature algorithms are described in detail. The feature operator circularity (*s*) calculates the similarity of the input region of the given image with a circle. If *F* is the area of the image region and *max* is the maximum distance from the center point to all contour pixels, the shape factor *s* is calculated as [11]:

$$s = \frac{F}{(\max^2 \cdot \pi)} \tag{1}$$

If the region is a circle then the shape factor *s* is calculated to the value 1. If the region has another shape as a circle, *s* is smaller than 1.

Another feature operator is the entropy and anisotropy coefficient of the image, defined as [11]:

$$entropy = -\sum_{0}^{255} rel[i] * ld(rel[i]) \,. \tag{2}$$

$$anisotropy = \frac{\sum_{0}^{k} rel[i] * ld(rel[i])}{entropy} \,. \tag{3}$$

The used parameters are:

- *rel*[*i*] – histogram of relative gray value frequencies,
- *i* – gray value of input image (0..255) and
- *k* – the smallest possible gray value.

The specific feature vectors consist of 188 feature values and provide the basis of the classifier training.

2.2. Best differentiating features - feature selection

The performance of several classification algorithms decrease by using redundant and irrelevant features in feature vector. In addition the calculation needs a lot of time for an excessive feature vector with non-informative features. But the aim of the investigations is a real-time recognition system of building materials.

For example the time a Support Vector Machine (SVM) needs for classification depends linearly on the dimension of the feature vector and number of support vectors. It is not expedient to calculate irrelevant features. In summary the aim is the reduction of the dimension of feature vector.

Filter selection methods are independent of any classifiers [8]. They filter out features with low discriminatory power. These methods are based on the performance evaluation metric calculated directly from the given dataset [8]. In contrast to wrapper methods, filter methods are normally not computationally intensive.

Wrapper methods are very time-consuming with using a complex training algorithm like the SVM. This fact rests in estimating the discriminatory power of features by calculation the recognition accuracy of each feature selection (feature subset).

So filter selection methods were used especially the Info-Gain-Attribute-Evaluator and Chi-Squared-Attribute-Evaluator.

The Information Gain measure is based on the entropy [7]. The entropy for the class distribution *C* is evaluated as:

$$H(C) = \sum_{c \in C} p(c) \cdot \log_2 \cdot p(c) \,. \tag{4}$$

The conditional entropy for class *c* and feature *F* is evaluated in the form of:

$$H(C \mid F) = \sum_{f \in F} p(f) \sum_{c \in C} p(c \mid f) \cdot \log_2 \cdot p(c \mid f) \,. \tag{5}$$

From this it follows that the Info Gain for the specific feature *i* is evaluated in the form of:

$$IG_i = H(C) - H(C \mid F_i) \,. \tag{6}$$

As a result a score is calculated for each single feature. A statement of the specific discriminatory power is represented by this score.

The Chi-Squared filter method estimates the distributional properties of a statistical basic population in consideration of a specific distributional property [8]. Setino et al. [9] found out, that the discretization is an appropriated instrument for selection of numerical features.

The discretization is carried out by using the Chi-Squared statistic. The Chi-Squared value has to be determined as the test statistic for a significance test.

Table 2 shows a selection of the best ranked texture and contour features by using Info Gain method. As a result the colour features and in particular texture features are the best discriminating features for the given problem far ahead of contour features. The most useful texture features can be calculated in the H- and S-channel of the HSI colour space. Independent of the two selection methods, nearly the same ranking list was calculated with similar ranking positions. The ranking list of features was used to build up different feature subsets with the number of best ranked features. This means, when the number of used best ranked features equals 16, the 16 best ranked features of the ranking list were used to build up a feature subset. Its discriminatory power is specify by determination the classification accuracy of a trained and tested classifier. Finally there is a statement given, how many best ranked features are needed for reaching a good classification performance.

Table 2. Selection of the best Info-Gain ranked features

	feature name	channel	description [8]	ranking pos.
colour and texture features	Fit_Surface_2nd_order_Zeta	S	calculates the gray value moments and the parameters of the approximation of the gray values by a second order surface, zeta is the most useful parameter	1
	Cooc_Feat_Energy	H	calculates the gray value feature energy from Cooc-matrix, energy is a measure of homogeneity of the gray values	2
	Fuzzy_Entropy1	S	calculates the fuzzy entropy of a fuzzy set, measure of how well the image approximates a white or black image	3
	Fit_Surface_2nd_order_Zeta	H	calculates the gray value moments and the parameters of the approximation of the gray values by a second order surface, zeta is the most useful parameter	4
	Moments_Gray_Plane_Mean1	S	calculates the gray value moments and the parameters of the approximation of the gray values by a plane	5
	Intensity_Mean	S	calculates the mean and the deviation of the gray values in the input image within Regions	6
	Fit_Surface_1st_order_Gamma	S	approximates the gray values by a first order surface, gamma is the most useful parameter	7
	Entropy	H	calculates the entropy of gray values in the image region	8
	max gray	I	creates the histogram of the absolute frequencies of the gray values within Regions in the input image Image and the largest gray value	9
	Fuzzy_Entropy1	H	calculates the fuzzy entropy of a fuzzy set, measure of how well the image approximates a white or black image	10
contour features	runlength_features_Kfactor	B	runlength features calculates for every input region the number of runs necessary for storing this region with the aid of runlength coding, „K-factor" indicates how much the number of runs differs from the ideal of the square in which this value is 1.0.	61
	runlength_features_Lfactor	B	L-factor indicates the mean number of runs for each line index occurring in the region	62
	Convexity	B	calculates the convexity of each input region	67
	runlength_features_MeanLength	B	MeanLength indicates the mean length of the runs	68

2.3. Supervised machine learning - used classifiers

There are a plurality of classifiers in the field of supervised learning. In the theory of machine learning Naive Bayes classifier, decision trees, k-nearest-neighbour algorithms, neural networks and SVM are common classifiers [12]. An overview of supervised learning algorithms is shown in figure 3 [13].

Different classification algorithms of the machine learning library Weka [14] were tested after feature extraction, such as LibSVM, Random Forest, k-nearest neighbour, Naive Bayes and J48 by using a 10-fold cross validation.

Especially the classification with SVM needs an optimal parameter selection for a good classification performance. First an introduction to the characteristic of SVM is given in this paper. The SVM was introduced by [15] and is mentioned as being one of the most powerful classifiers today. It is derived from the statistical learning theory [16]. The algorithm is motivated by the structural risk minimization, which says that not only the training error but also the complexity of the model influences its generalization ability. The SVM was designed to solve binary classification problems but there are different strategies to solve multi-class problems, too. The SVM executes a non-linear projection of data in a higher dimensional feature space. The classes are separable in a linear way in this higher dimensional feature space. During the training process an optimal hyperplane is constructed. Optimal means that it leaves a maximal margin between the hyperplane and the closest training point on both sides.

The kernel function $k(x,x_i)$ extends the linear discriminant SVM to a nonlinear machine. The given decision function is [8]:

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{m} \alpha_i y_i k(x, x_i) + b\right). \qquad (7)$$

The radial basis function kernel (rbf) is one of the popular kernel functions and is defined as:

$$k(x, x') = e^{-\gamma \|x - x'\|^2} \qquad (8)$$

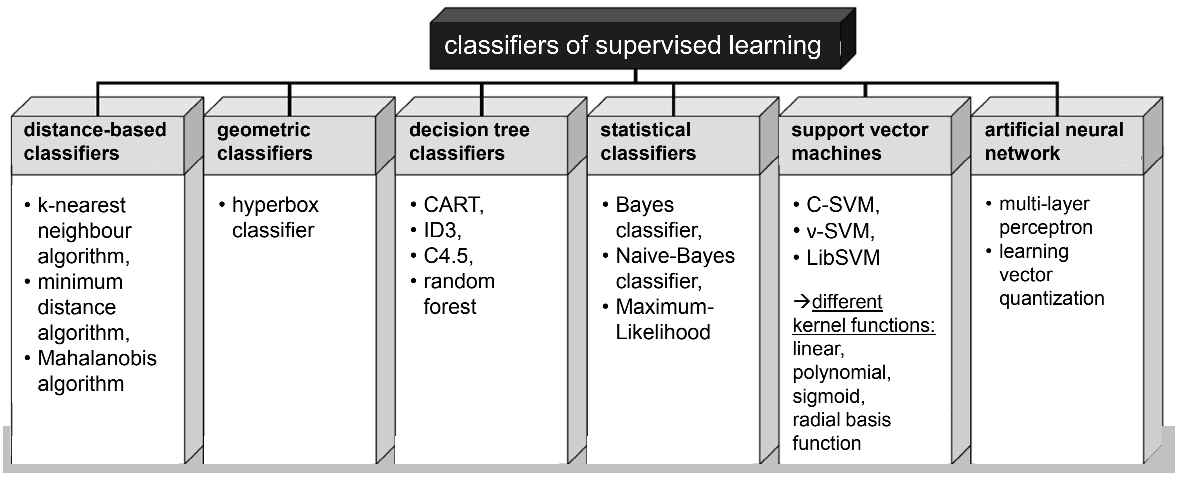

Fig. 3. Classifiers of supervised learning [13]

In the investigations the LibSVM [17] of Weka was tested with the rbf-kernel and different parameter modifications for the cost parameter C of C-SVC and the parameter γ in kernel function. The optimal chosen parameters C and γ are unknown before starting an optimization process. But they are very important for reaching the best classification performance for the given problem. The goal is to accurately predict unknown data, i. e. data which were not used for training. So the model selection or parameter search have to be done in a computationally efficient way.

The optimization was done on the 188 dimensional feature space. A 3-fold cross validation was used for training the LibSVM with different parameter selections to prevent the overfitting problem. The search was done by using a grid search for the different steps of parameter modification. In a first step the exponentially growing sequences of C and γ as practical method was used to identify good parameters (see [18]). In the second step a fine grid search was used with an equidistant increment on the identified "better" region on the grid. Fig. 4 shows the results of the optimization process.

The parameter γ has a considerably higher influence on the total recognition rates than the parameter C. The highest accuracy with 98,2 % was reached by using $\gamma = 0{,}05$ and $C = 100$. The best compromise is to use $\gamma = 0{,}15$ and $C = 30$ to reach a total recognition rate of 98,1 %, because it is better to use a lower C as penalty parameter of the error term.

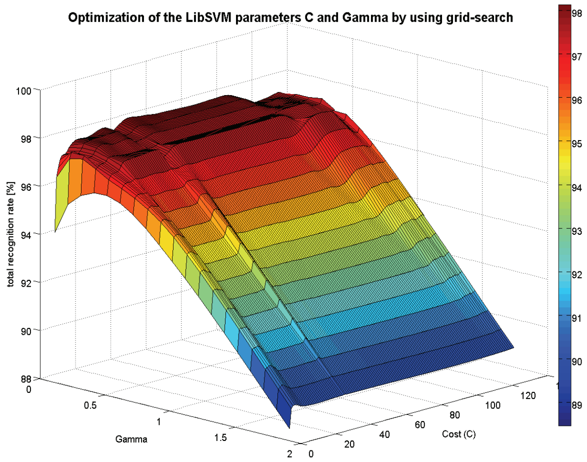

Fig. 4. Surface plot of parameter optimization of LibSVM

In the comparison of different classifiers (see figure 5) the compromise of parameter selection was used for LibSVM.

The choice of an optimal classification algorithm is an important task of investigations in pattern classification. So some classification algorithms were tested for the given dataset by using several feature subsets, which were found with the feature selection method Info Gain. The results, total recognition rates (TRR) and recognition rates per class (RR), are shown in figure 5 and figure 6.

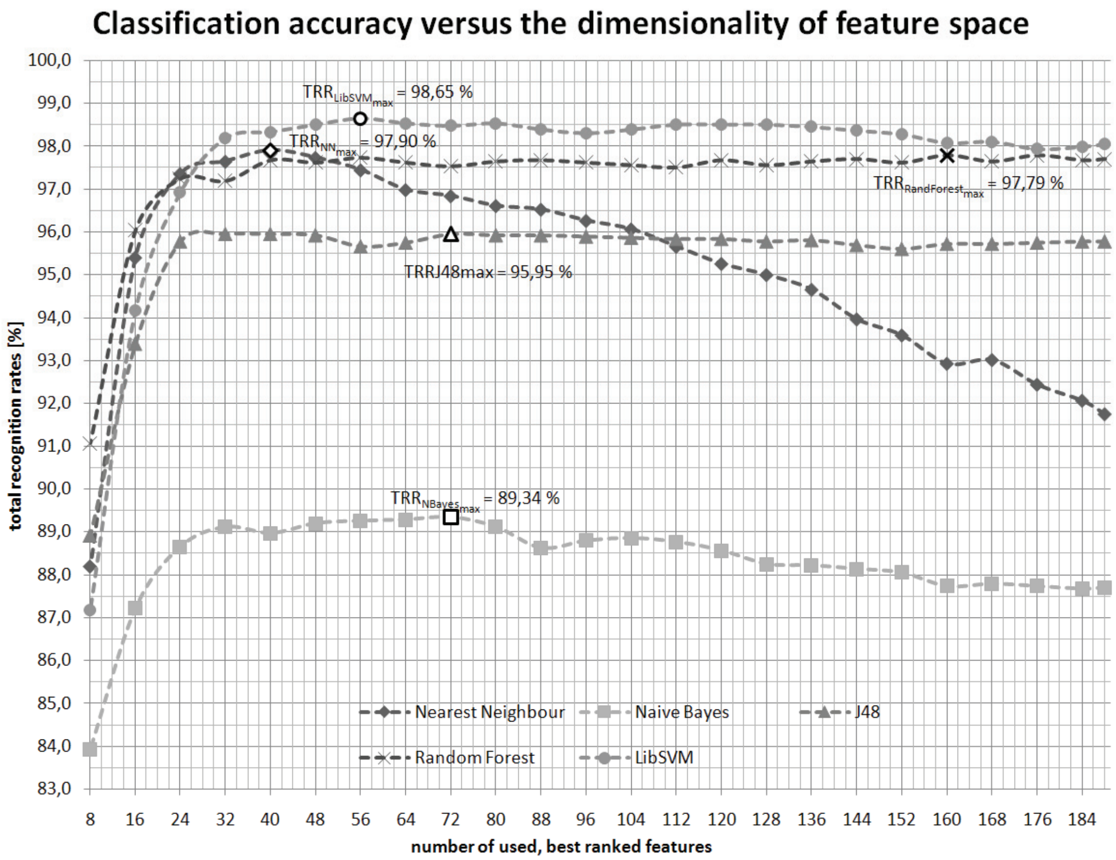

Fig. 5. Classification performance of different classifiers versus dimensionality of feature space by using the number of best ranked features

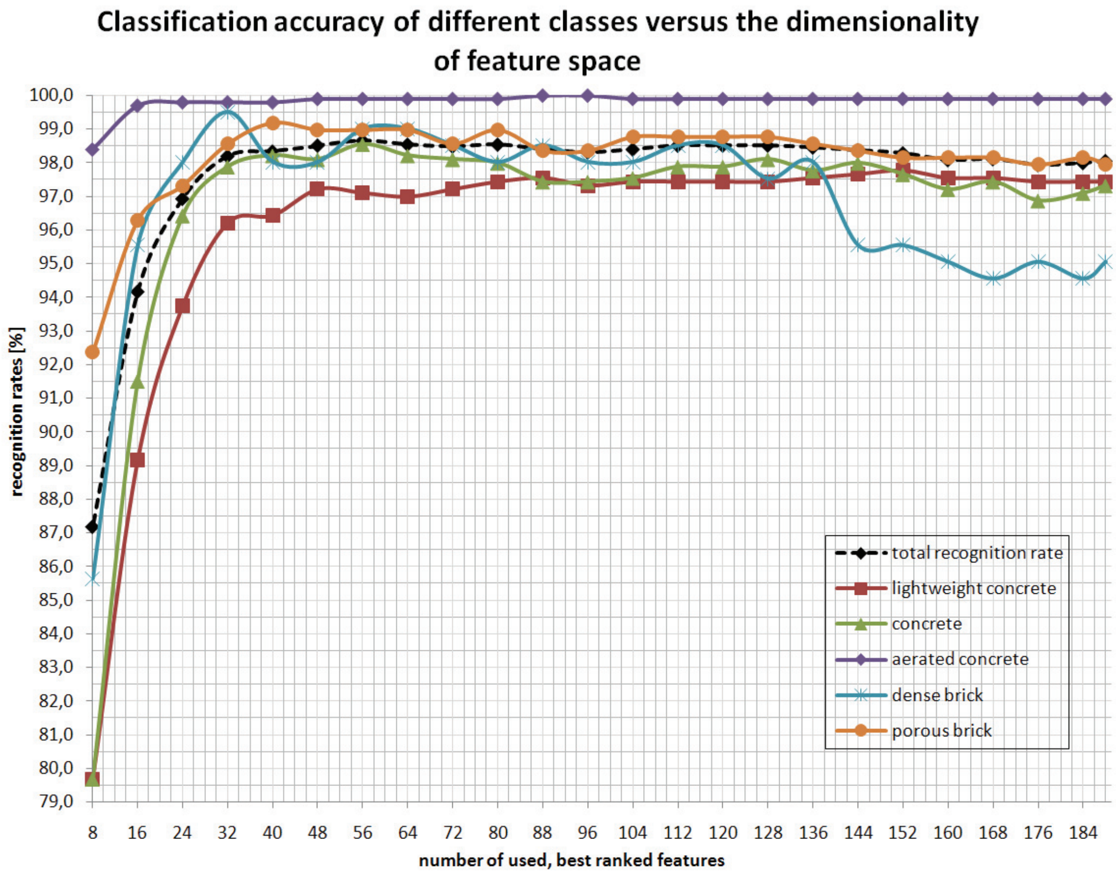

Fig. 6. SVM performance versus dimensionality of feature space by using the number of best ranked features

The investigation shows that the LibSVM (with TRR = 98,7 %) and the Random Forest classifier (with TRR = 97,8 %) are the best classifiers for the recognition of building materials out of the plurality of tested classifiers.

In the experiments Naive Bayes reached the lowest total recognition rate (89,3 %). Better results could be observed for J48 classifier (96,0 %) and nearest neighbour (97,9 %).

In addition to this, only the classifiers nearest neighbour and Naive Bayes show overfitting by using an excessively complex model. They still have too many features relative to the number of given data samples.

The classwise reached individual recognition rates of the best classifier, the parameter optimized LibSVM, are explained in the following (see figure 6). The individual recognition rates were calculated by using the different feature subsets. For all classes very good individual recognition rates over 96 % was reached. The best recognition rates were reached for the class aereted concrete (100 %). Similar good recognition rates were reached for the other 4 classes in the range of 97,8 % to 99,5 %. This performance seems to be very high but the level of complexity of the given problem has to be kept in mind. At first only 5 classes were used out of the plurality of building classes. If more classes are used in further investigations, the overall classification performance will be decrease – approximately 10 percent less than the amount before.

4. CONCLUSION

In this approach a part of relevant building material classes were used for testing different feature subsets and classification algorithms for the given recognition problem.

Two different filter selection methods were used for the detection of redundant and irrelevant features in feature vector. As the result, the colour features and in particular the texture features in the H- and S-channel are the best discriminating features for the given problem far ahead of contour features.

After this the ranking list of features was used to build up different feature subsets with the number of best ranked features. Then different classification algorithms of the machine learning library Weka were tested, such as LibSVM, Random Forest, k-nearest neighbour, J48 and Naive Bayes by using a 10-fold cross validation.

The cost parameter C and the kernel parameter gamma of LibSVM were optimized to reach best classification performances and prevent overfitting. The optimization was done on the 188 dimensional feature space by using grid search. Finally it is pointed out that the parameter γ has a considerably higher influence on the performance than the parameter C.

In addition the approaches demonstrated the SVM and Random Forest as the best classification algorithms for this recognition task. The parameter optimized LibSVM achieved a total recognition rate of 98,7 % and the Random Forest classifier of 97,8 % for the given dataset. This agrees with the fact, that SVM and Random Forest are two of the most efficient classifiers today. This fact could also be demonstrated in previous investigations [13].

In future investigations the dataset has to be optimized. The dataset has to be extended for other relevant building classes. The characteristic object features of each class and their statistical distribution have to be specified.

REFERENCES

[1] *Monitoring Report by ARGE (Monitoring-Bericht der ARGE)* Kreislaufwirtschaftsträger Bau, Berlin, 2004.

[2] C. F. Hendriks, *The building cycle,* Aeneas, Netherlands, 2000.

[3] A. Mueller, *Closed loop of concrete rubble? Lecture notes,* Bauhaus-University of Weimar. available in: www.uni-weimar.de/Bauing/aufber/Lehre/ Gastvorlesung/Barcelona/Lecture07.pdf, 2007.

[4] *RECDEMO: Complete utilization of the sand fraction from demolition waste recycling,* Laymann report. EU LIFE, 2004.

[5] M. Dehler, *Optical Sorting of Mineral Commodities* (*Optische Sortierung von mineralischen Rohstoffen)*, Onlinefassung, 2003.

[6] COMMODAS, *We sort it out*, Catalog. 7 p. http://www.commodas.de, 2007.

[7] T. M. Mitchell, *Machine Learning*. The Mc-Graw-Hill Companies Inc, 1997.

[8] I. Guyon, A. Elisseeff, *Feature extraction: foundations and applications*, Berlin, Springer Verlag, 2006.

[9] R. Setino, H. Liu, *Chi²: Feature selection and discretization of numeric attributes*, In: Proceedings of the Seventh IEEE International Conference on Tools with Artificial Intelligence, 1995.

[10] *QI Inspector: ergonomic vision and installation workstation with colour camera (QI Inspector: Ergonomischer Sicht- und Montagearbeitsplatz mit Farbkamera)*, available in: http://www.stz-ilmenau.de/products/, 2011.

[11] Halcon, Reference Manual - Halcon 8.0.3, MVTec Software GmbH, www.mvtec.com, 2009.

[12] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification*, A Wiley Publication, 2001.

[13] K. Anding, *Automatic Quality Assurance of Wheat by Using Supervised Machine Learning Methods in Image Processing (Automatisierte Qualitätssicherung von Getreide mit überwachten Lernverfahren in der Bildverarbeitung)*, dissertation, Ilmenau University of Technology, 2010.

[14] M. Hall, F. Eibe, G. Holmes, B. Pfahringer, P. Reutemann, I. H. Witten, *The WEKA Data Mining Software: An Update*, SIGKDD Explorations, Vol. 11(1), 2009.

[15] B. Boser, I. Guyon, V. Vapnik, *A Training Algorithm for Optimal Margin Classifiers*, Proceedings of the Fifth Annual Workshop on Computational Learning, 1992.

[16] V. Vapnik, *The Nature of Statistical Learning Theory,* New York, Springer Verlag, 1995.

[17] Chih-Chung Chang, Chih-Jen Lin, *LIBSVM : a library for support vector machines*. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software available at http://www.csie.ntu.edu.tw/ ~cjlin/libsvm.

[18] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, *A Practical Guide to Support Vector Classification*, Department of Computer Science National Taiwan University, http://www.csie.ntu.edu.tw/~cjlin, 2010.

Author(s):
Dr. Katharina Anding, Ilmenau University of Technology, Faculty of Mechanical Engineering, Department of Quality Assurance and Industrial Image Processing, Gustav-Kirchhoff-Platz 2, 98693 Ilmenau, Germany,
phone: +49 3677/693964, fax: +49 3677/693823, e-mail: katharina.anding@tu-ilmenau.de