

Framework for developing image-based dirt particle classifiers for dry pulp sheets

Nataliya Strokina · Aki Mankki · Tuomas Eerola ·
Lasse Lensu · Jari Käyhkö · Heikki Kälviäinen

Received: 29 October 2011 / Revised: 7 December 2012 / Accepted: 15 January 2013 / Published online: 6 March 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract One important aspect of assessing the quality in pulp and papermaking is dirt particle counting and classification. Knowing the number and types of dirt particles present in pulp is useful for detecting problems in the production process as early as possible and for fixing them. Since manual quality control is a time-consuming and laborious task, the problem calls for an automated solution using machine vision techniques. However, the ground truth required to train an automated system is difficult to ascertain, since all of the dirt particles should be manually segmented and classified based on image information. This paper proposes a framework for developing and tuning dirt particle detection and classification systems. To avoid manual annotation, dry pulp sheets with a single dirt type in each were exploited to generate semisynthetic images with the ground truth information. To classify the dirt particles, a set of features were com-

puted for each image segment. Sequential feature selection was employed to determine a close-to-optimal set of features to be used in classification. The framework was tested both with semisynthetically generated images based on real pulp sheets and with independent original real pulp sheets without any generation. The results of the experiments show that the semisynthetic procedure does not significantly change the properties of images and has little effect on the particle segmentation. The feature selection proved to be important when the number of dirt classes changes since it allows to improve the classification results. Using the standard classification methods, it is possible to obtain satisfactory results, although the methods modeling the data, such as the Bayesian classifier using the Gaussian Mixture Model, show better performance.

Keywords Image processing and analysis · Machine vision · Ground truth generation · Particle segmentation · Dirt particle classification · Feature selection · Pulping · Papermaking

N. Strokina (✉) · T. Eerola · L. Lensu · H. Kälviäinen
Machine Vision and Pattern Recognition Laboratory (MVPR),
Department of Information Technology, Faculty of Industrial
Management, Lappeenranta University of Technology (LUT),
P.O. Box 20, 53851 Lappeenranta, Finland
e-mail: nataliya.strokina@lut.fi
URL: <http://www2.it.lut.fi/mvpr>

T. Eerola
e-mail: Tuomas.Eerola@lut.fi

L. Lensu
e-mail: Lasse.Lensu@lut.fi

H. Kälviäinen
e-mail: Heikki.Kalviainen@lut.fi

A. Mankki · J. Käyhkö
Fiber Laboratory, Mikkeli University of Applied Sciences,
Vipusenkatu 10, 57200 Savonlinna, Finland
e-mail: Aki.Mankki@mamk.fi

J. Käyhkö
e-mail: Jari.Kayhko@mamk.fi

1 Introduction

To optimize its production processes, the pulp- and papermaking industry is searching for intelligent solutions to assess and control product quality. Optimization, in this context, can be defined as building resource-efficient and environmentally sound production with known quality, using less raw materials, water, and energy. This optimization is not easy since the pulp- and papermaking process consists of a large number of stages where the treatment of the material can have a significant impact on the properties of the final product. Therefore, it is important to know how the important characteristics of the product are formed at each stage.

The relevant properties of pulp and paper are traditionally evaluated in a paper laboratory, where the procedures follow the standards related to quality control and the properties of the raw material or end-products are measured from samples taken from the process [13]. Since traditional quality control is time-consuming and does not allow the direct control of the production process, the industry is interested in transferring the laboratory measurements to the inline process. In this scenario, the measurements will be made in real-time directly of the material in the process to reduce the delay in obtaining quantitative quality information and enable even real-time process control. The amount of the produced pulp and paper at the industrial scale is considerable, which means that the inline solution offers significant benefits for the industry, allowing it to adjust the process according to the information obtained from the process measurement. This would considerably reduce the risk of producing a large amount of products with undesired properties.

The detection and classification of dirt is an important part of pulp and paper quality assessment. This is because dirt causes undesired surface properties in subsequent processing, negatively affects the surface appearance, and can impair the printability of paper. Commonly, the detected unwanted matter in pulp or paper is an indicator of less-than-optimal operation or a malfunction of the process. Therefore, knowledge about the dirt type can help to identify the problematic stage and to perform maintenance at that stage. To detect such unwanted matter, methods have been developed to segment and count the dirt particles. The concentration of dirt in pulp can be evaluated from sample sheets prepared from pulp suspension, by screening methods or when the dried pulp sheets are prepared for shipping. In this work the dirt particles are detected and classified in the dried pulp sheets.

Dirt detection and counting have been studied earlier, but the problem of dirt particle classification was not addressed. For example, Fastenau et al. [5] present a laser system for dirt counting on the industrial scale. The paper gives the motivation for the automation of the dirt counting process, explaining the difficulties of the manual procedure. Based on the shape of the obtained signal, the system was capable to perform dirt particle categorization by the size of particles. Sutman [23] presents a method for measuring the testing precision. The effect of a sample size on dirt count test precision was not well understood and it was the motivation for the research. Rosenberger [18] showed that the threshold for dirt counting should be selected automatically as well to be able to adapt for different lighting conditions as well as paper and dirt particle properties.

Juntunen et al. [10] introduce an automated analysis system for colored ink particles in recycled pulp. The samples were prepared with a known percentage of ink. A microscope with an attached color video-camera was

used to image the samples. Thresholding was performed separately for three HSI channels, followed by the connectivity analysis. The system allowed to obtain the dirt counts and to measure the size distribution of the particles. Since the ground truth did not contain the information on the location of the particles, there was no opportunity to judge about the spacial distribution of the dirt.

In Ref. [15], the “Pulp Automated Visual Inspection System”, which segments and counts dirt particles, as areas in an image with an intensity lower than a certain threshold. Another example InsPulp, an on-line visual inspection system, is explained in Ref. [1]. The paper is imaged by a CCD line and the dirt is segmented using a local dynamic threshold, which allows the system to segment and detect the impurities in pulp with a low error rate. These methods only count the dirt particles and do not address the more challenging problem of dirt particle classification.

The industrial dirt counter system by VERITY IA [25] can divide the particles into a few groups based on their shape, but is still not able to identify the specific dirt types. The accurate classification of particles would be a great benefit. Savings in chemical and energy consumption could be attained by adjusting bleaching and screening, the aim of which is to eliminate the impurities in the material. In a production problem situation, fast and precise information on the type of particles present in the process can reveal the source of the problem, and the process can be adjusted accordingly.

One of the major problems in developing an automated dirt particle classification system is the collection of ground truth data essential for training of a supervised system. To obtain the ground truth, the exact location and type of each dirt particle need to be given. The identification of specific dirt particles can be a very difficult task even for experts, and the large amount of data required makes collecting the ground truth a very laborious and time-consuming process. In some systems, the difficulties with the performance evaluation are mentioned. For example, in Ref. [15] there was no opportunity to compare the results with manually segmented particles. In Ref. [1], it is also shown that an inspection by humans may be subjective: the number of dirt particles detected by different inspectors was different.

In their previous publications, the authors have presented a solution for generating the ground truth [22] and conducted initial experiments concerning the dirt classification problem [21]. In this work, the previous studies are combined and extended to build a full framework for developing dirt particle detection and classification systems. The framework was designed using the provided laboratory paper sheets, but it can be transferred to the industrial scale. The expertise of the paper laboratory personnel is not used to provide the ground truth manually but for

carrying out their most important task: produce dirt and pulp as clean as possible. This paper introduces a method for generating the ground truth and methods for dirt particle classification. The framework was tested both with semi-synthetically generated images based on real pulp sheets and with independent original real pulp sheets without any generation.

The paper is structured as follows: Sect. 2 provides a general scheme of the framework for developing dirt classification systems and demonstrates the initial data. Section 3 describes the solution for generating the ground truth. Section 4 handles the question of feature extraction, evaluation, and dirt particle classification. Section 5 discusses the results of the experiments, and the conclusion is drawn in Sect. 6.

2 Framework for developing dirt particle classification

2.1 Data

The initial data are represented by scanned images of dry pulp sample sheets. The samples were provided by pulp and papermaking experts from the Fiber Laboratory, a collaborating unit with expertise in pulping and papermaking.

The samples used in this study consisted of three different pulp types, and four different types of dirt particles. The pulp types were bleached hardwood, bleached softwood, and softwood pulp after the second chlorine dioxide bleaching stage (D_1), the color of which is not completely white (see Fig. 1). Thus, more variation for the background was gained. Despite the fact that the number of different pulp and dirt types is low

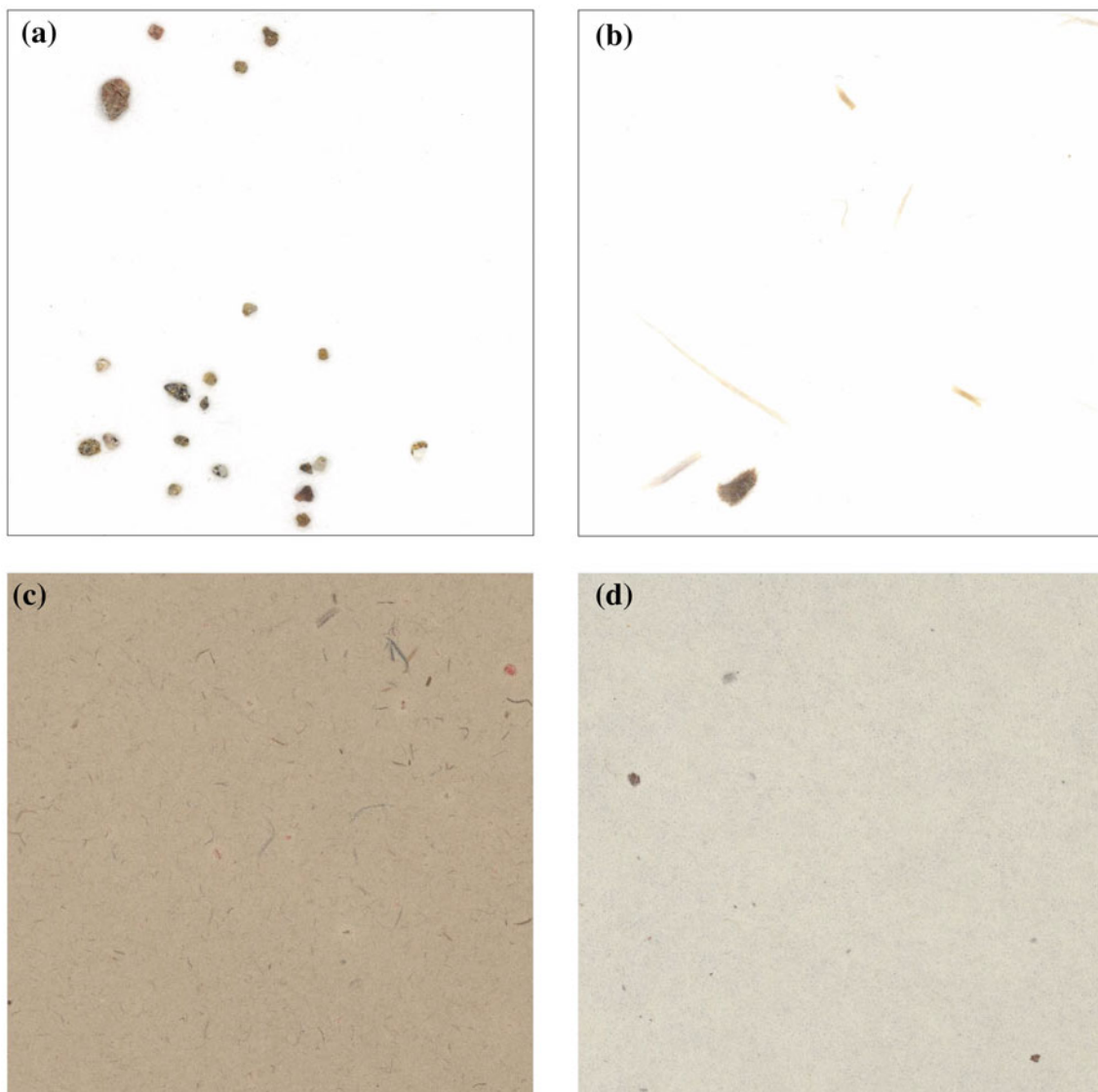


Fig. 1 Pulp sheet images: **a** bleached pulp with sand, **b** bleached pulp with shive, **c** D_1 pulp with bark, **d** stock pulp with plastic



Fig. 2 Clearly different examples of the dirt particles: **a** bark, **b** plastic, **c** sand, **d** shive

and cannot be considered to represent the full variation of pulp in the industry, the sample set was sufficient to develop the framework.

Four common types of dirt particles were selected based on the literature [2,9] and expert knowledge: shives, bark, plastic, and sand (see Fig. 2). The dirt particles were either prepared or separated from the pulp in a paper laboratory. The shives were separated from reject pulp from brown stock screening. The bark particles were created by disintegrating pine bark mixed with water in a disintegrator. A plastic canister was grinded to create excess plastic particles. The natural sand was washed to get rid of extra particles and dust. A small amount of sand was also obtained as reject pulp was washed.

Sample sheets of the three different pulp types were prepared according to the ISO 5269-1 standard [17]. The amount of pulp equivalent to 1.63 g of dry pulp and an adequate amount of a single type of dirt particles were mixed before sheet forming. Consequently, the prepared sample sheets contained an amount of fibers equivalent to a standard 60 g/m² sample sheet and an adequate amount of dirt particles. The amount of dirt particles was controlled so that there would be more than 20 particles per sample sheet, but not too many to avoid significant overlapping of the parti-

cles. All of the three pulp types, one at a time, were mixed with one of the four types of dirt particles. Five sample sheets per test point and also one sample containing sand separated from industrial pulp were prepared. As a result, the sample set consisted of 61 sample sheets.

To image the samples, the pulp sheets were scanned with a professional Microtek ArtixScan 2500f scanner with 1250 dpi (A4) resolution, true 42 bit RGB colors, and under reflected light. According to the ISO 5350-1:2006 standard [16], the minimum size of particles that are to be detected is 0.04 mm². The physical resolution was around 0.0004 mm² per pixel, which means that the smallest dirt particle occupies, approximately, 100 pixels.

2.2 Workflow description

The basic idea of the proposed framework is illustrated in Fig. 3. The initial images with a single dirt type in each are used to produce the semisynthesized images. From each of the images, segmented dirt particles are collected to create a database of dirt particle images which are to be scattered on a generated background. The method to produce the semisynthetic background is discussed in Sect. 3. The actual ground truth is represented by a binary mask, containing the exact location of each particle and its type. After the semisynthesized images are obtained, one can segment and classify dirt, which can be evaluated using the semisynthetic ground truth.

The semisynthetic images are processed using the following Algorithm 1 where a set of features are extracted from the segmented particles and the close-to-optimal feature set is determined to be used in further classification.

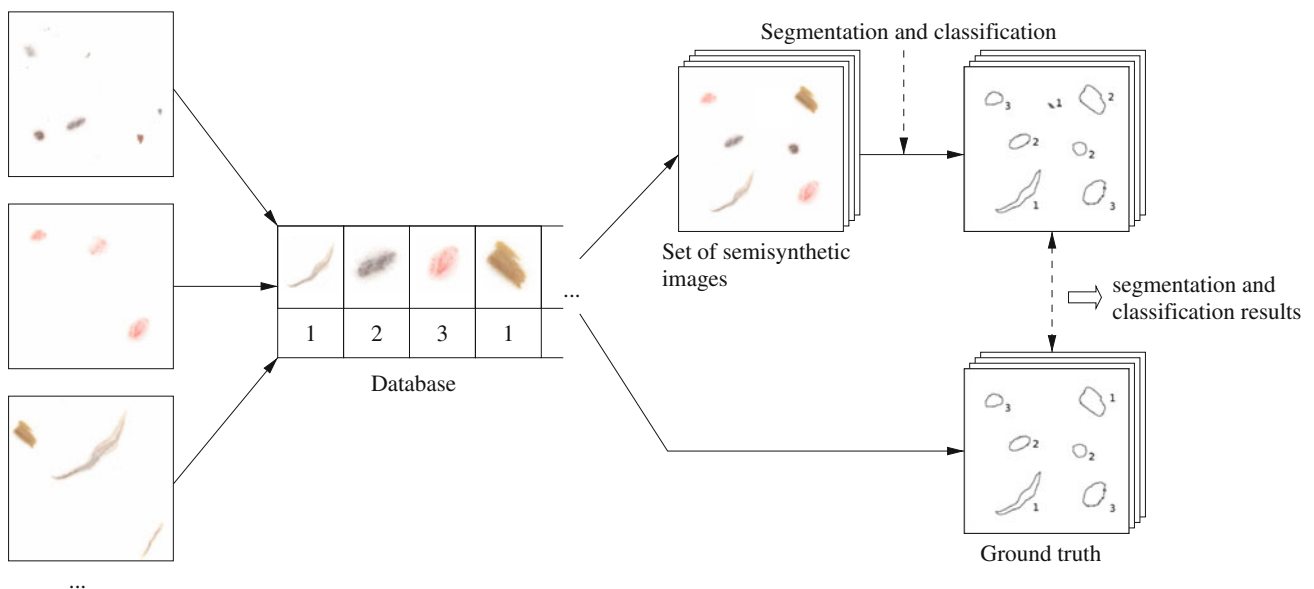


Fig. 3 Framework for dirt particle segmentation and classification

Algorithm 1 General workflow for dirt particle segmentation and classification

- 1: Use a segmentation method such as Kittler thresholding [11] to produce the dirt image segments (Sect. 3.1).
- 2: Manually validate the segmented dirt.
- 3: Extract features from the segments (Sect. 4.1).
- 4: Determine a close-to-optimal feature set for classification (Sect. 4.1).
- 5: Train and test of different classifiers (Sect. 4.2).
- 6: Evaluate the performance of the classifiers (Sect. 5).

3 Semisynthetic ground truth generation

In any classification system, it is important to have reference data to allow the evaluation of the classification result. As already mentioned, the ground truth can be, in several cases, laborious or even impossible to obtain. Therefore, there is a need to produce the reference data synthetically. Manual dirt annotation can be substituted by the semisynthetic procedure that is described in this section. The semisynthetic ground truth generation consists of the following stages: (1) dirt particle segmentation and database generation, (2) background generation to fill the holes left by removed dirt particles, and (3) random scattering of the dirt particle images and creation of the corresponding ground truth image.

3.1 Segmenting images of sheets with dirt

The first step of semisynthetic ground truth generation is to create a database of dirt particles. To accomplish this, dirt particles need to be segmented from the initial images. According to the survey [19], there exists a number of methods to segment foreground objects (dirt particles) from the background, but none of them can be universally used for segmentation problems. In this study, the Kittler thresholding method [11] is used. The choice is based on previous experimental studies on detecting small particles from fiber-based surfaces [3, 7]. A global thresholding approach should be adequate for segmentation purposes in this case since the image backgrounds are rather homogeneous.




Gray-scale images are considered for segmentation, because the segmentation method is based on the intensities. The images are divided into the foreground, which consists of the dirt particles, and the background. The foreground and background are modeled as a mixture of two Gaussians [11]. The threshold can be calculated by optimizing the cost function based on the Bayesian classification rule.

The segmented particle images are used to create the database. Table 1 introduces the scheme of the database. An image with a bounding box for each of segmented particle is stored into the database, as well as its area and type.

3.2 Background generation

The segmented dirt is removed from the images so that the area occupied by the dirt particles is substituted by

Table 1 The scheme for the dirt particle database

Particle 1	Particles 2	Particle 3	...
			...
Type 1	Type 2	Type 3	...

white pixels. These holes in the image are then filled by using a Markov Random Field (MRF) method [26], which is one of the common approaches for texture synthesis, e.g., in [12, 20]. In the concept of MRFs [28], a random field is composed as pixel by pixel. The probability of each pixel to have a certain intensity depends on the intensity or color of its neighboring pixels. Before background generation, the user commonly specifies a sample region from the image manually, such that it contains no dirt particles and the background is reasonably uniform. This sample image is used as the ideal background for the area filling. In Algorithm 2, the probability is not calculated explicitly to reduce the computational load. The number of iterations can be restricted, for example, by a threshold dependent on the standard deviation of color which should be close to the original background.

Algorithm 2 Background generation

- 1: Construct the neighborhood of each pixel in the sample image.
- 2: Repeat the followings steps until a satisfactory result defined by the threshold is obtained.
 - a. Construct the neighborhood of the current pixel from the original image.
 - b. Find the best matching neighborhood from the sample image.
 - c. Substitute the current pixel by the pixel with the best matching neighborhood.

There are several factors to consider in this algorithm: (1) how to determine the neighborhood?, (2) how many iterations should be carried out?, and (3) what the best matching neighborhood is? In the implemented version of the algorithm, the neighborhood is determined by the user, as well as the number of iterations. The best match is found by computing the distance between two neighborhoods and choosing the one with the shortest Euclidean distance.

3.3 Inclusion of dirt particles

After the dirt database has been created and the holes left by the removed dirt particles are filled, the dirt particles can be embedded into the background images to create semisynthetic images. During this stage, a specified number of dirt particles is spread over the uniform synthesized background. According to Ref. [27], the particles should be placed randomly, and the implementation includes an option of random

rotation of the particles. To prevent dirt particle overlaps, a binary mask is used to store the information on occupied areas in the image. In Algorithm 3, the place for a new particle is selected from the unoccupied area.

Algorithm 3 Dirt particle placing

- 1: Initialize a binary mask to zeros.
 - 2: Repeat the following steps to reach the desired count of dirt particles.
 - a. Randomly select a dirt particle from those that have not been placed yet.
 - b. Find the place for the particle in the binary mask not occupied by any other particle.
 - c. Place the dirt particle on the background.
-

One aspect that should be taken into account is controlling the amount of dirt in an image. It can be controlled by the percentage of the dirt covering the surface. To form the list of dirt particles to be positioned on the image, the user should specify what type of dirt is needed to be added onto the image and its proportion.

As its result, Algorithm 3 outputs the image with a uniform background and dirt particles placed on it randomly. In addition, information about the dirt particles is stored in the form of a labeled binary mask (the ground truth).

In some cases it might be useful to consider normalization of the colors [22]. In this study, the color normalization was not used.

4 Dirt features and their use in classification

The data used in this study cannot be considered to represent the full variation of pulp nor all characteristics of dirt types. To make the system stable when a new type of dirt is introduced or to adapt the system for completely new dirt types, a tool is required to tune the classification process. This calls for feature evaluation, which aims at estimating the feature importance and makes it possible to determine the set of features that describe the available data in the most efficient way.

4.1 Feature extraction and evaluation

The dirt features used in this work can be divided into two categories: geometric features and color features. The geometric features include characteristics related to the shape, form, and uniformity of dirt particles. To complement them, the color features include, for example, mean color, variation of color, and intensity. The features are presented in Table 2.

For each dirt particle, a bounding box is determined which is the smallest rectangle enclosing the dirt particle. The solidity specifies the proportion of the pixels in the convex hull

Table 2 The feature set

Maximum diameter	Extent	
Minimum diameter	Length	Eq. 1
Solidity	Width	Eq. 2
Eccentricity	Form factor	Eq. 3
Convex area	Roundness	Eq. 4
Perimeter	Aspect ratio	Eq. 5
Std of color	Elongation	Eq. 6
Mean intensity	Curl	Eq. 7
Mean color	Coarseness	Eq. 8
Area		

The equation numbers refer to the text

that belong to the region. Eccentricity specifies the eccentricity of the ellipse that contains the same second-moments as the region. The convex area is the number of pixels in the convex hull of a dirt particle. The extent specifies the ratio of pixels in the region to pixels in the total bounding box. The mean color and mean intensity are calculated as the mean hue value and the mean intensity over a dirt particle area. Std of color describes the standard deviation of color within the area of a dirt particle. The other geometric features are calculated according to the following formulas:

$$\text{Length} = 0.25 \cdot (\text{Perimeter} - \sqrt{|\text{Perimeter}^2 - 16 \cdot \text{Area}|}) \quad (1)$$

$$\text{Width} = \frac{\text{Area}}{\text{Length}} \quad (2)$$

$$\text{FormFactor} = \frac{4 \cdot \pi \cdot \text{Area}}{\text{Perimeter}^2} \quad (3)$$

$$\text{Roundness} = \frac{4 \cdot \text{Area}}{\pi \cdot \text{MaxDiameter}^2} \quad (4)$$

$$\text{AspectRatio} = \frac{\text{MaxDiameter}}{\text{MinDiameter}} \quad (5)$$

$$\text{Elongation} = \left| \frac{\text{Length}}{\text{Width}} \right| \quad (6)$$

$$\text{Curl} = \left| \frac{\text{MaxDiameter}}{\text{Length}} \right| \quad (7)$$

$$\text{Coarseness} = \frac{\text{Perimeter}^2}{4 \cdot \pi \cdot \text{Area}} \quad (8)$$

To determine the feature set to be used in classification, one should choose a feature evaluation procedure. The exhaustive search, i.e. evaluating all the possible combinations of features, would be computationally infeasible.

This is why the sequential feature selection [14] procedure is used instead. At each step, the method adds (forward selection) or removes (backward selection) a new feature to the feature set and calculates an objective function value, which is aimed to be minimized. Since the method moves only in

Table 3 The list of used classifiers

K-Nearest Neighborhood (k-NN)	[24]
Support vector machine (SVM)	[4]
Naive Bayesian classifier (Bayes)	[24]
Bayesian classifier using Gaussian mixture model (GMM)	[4]
Linear discriminant analysis (LDA)	[8]

one direction, adding or deleting features, it eventually evaluates only a subset of all possible feature combinations.

Therefore, the optimal feature set cannot be guaranteed but rather a close-to-optimal feature set is produced. In this work, the sequential feature selection was applied using the forward selection and the linear discriminative function [4]

$$g(\mathbf{x}) = \mathbf{w}^t \mathbf{x} + w_0 \tag{9}$$

was selected as the objective function. In Eq. 9 the weights \mathbf{w} in the linear combination of the features \mathbf{x} are optimized to minimize the linear discriminative function, taking into account the bias w_0 . This provides the information how well the combinations of features can be distinguished.

4.2 Classification methods

The classifiers used in the study are listed in Table 3.

State-of-the-art generic classification methods as well as the well-known structural approaches are used to avoid being related to specific data. k-NN is used with neighborhoods of 1, 3, and 5 samples. The goal of the Linear Discriminant Analysis (LDA) is to maximize the separability of data classes, defining the transform to the space where the current features might be distinguishable in the most efficient way.

The Gaussian Mixture Model (GMM) classifier is used with expectation maximization (GMMem) and Figueiredo-Jain (GMMfj) criteria [6]. With the Figueiredo-Jain criterion, the maximum number of 20 components were set and as a result each class was modeled with the most appropriate number of components, the maximum number being 7. For this reason, in the expectation maximization approach, seven components were used to model the data.

Support Vector Machine (SVM) was used with the radial basis function kernel.

5 Experiments and results

The experiments carried out in this research can be divided into three groups: (1) the evaluation of the quality of the semisynthetic images, (2) the estimation of the performance of the classification approaches, (3) the analysis of how the semisynthetic data affect the segmentation and classification results, and (4) segmentation and classification of dirt parti-

cles in the real pulp sheets based on the semisynthetic training data.

All of the experiments were carried out using the data presented in Sect. 2.1.

5.1 The semisynthetic data and the statistical evaluation of the generated background

Figure 4 demonstrates the steps of the semisynthetic background generation. The dirt particles are removed from the target image and the holes are filled according to the procedure presented in Sect. 3.2.

Figure 5 demonstrates the examples of the generated semisynthetic images (see Sect. 3.3). The semisynthetic images contain dirt particles of a predetermined amount, which is set by the user. It defines what percent of the total image area the dirt occupies. The ground truth is stored as a mask with labeled components, where the dirt particles of the same type have the same label. Figure 5b, d represents different types of dirt by different intensity values in the mask images.

To evaluate the quality of the generated background, the *F* test was performed to compare the variances of the generated and original parts of the background. The null hypothesis, i.e., the generated and original background have the same variance could not be rejected at the 2 % significance level, which means that the difference between the variances of the generated and original parts of the background is statistically insignificant.

5.2 The effect of the semisynthetic procedure on the segmentation

The semisynthetic procedure has an influence on the background intensity values as the particles are rotated and shifted, and therefore, it is important to know how the semisynthetic procedure affects the segmentation. In the classification, on the other hand, only the segmented dirt particles are involved. Therefore, the classification is not significantly affected by the semisynthetic data generation. For the 12 original images, 3 per each dirt class, the corresponding semisynthetic images were generated in order for the dirt from one image to remain only in that image, and no other dirt particles were used. In other words, the semisynthetic images contained only those dirt particles which were located in the corresponding original ones, but the particles were moved and rotated according to the semisynthetic procedure. Subsequently, the dirt particles were detected and counted in both original and generated images. The number of detected particles in the semisynthetic images was compared to the number of particles detected in the original ones, and the results are presented in Table 4. The table shows that the results differ insignificantly for the semisynthetic and original images. Yet

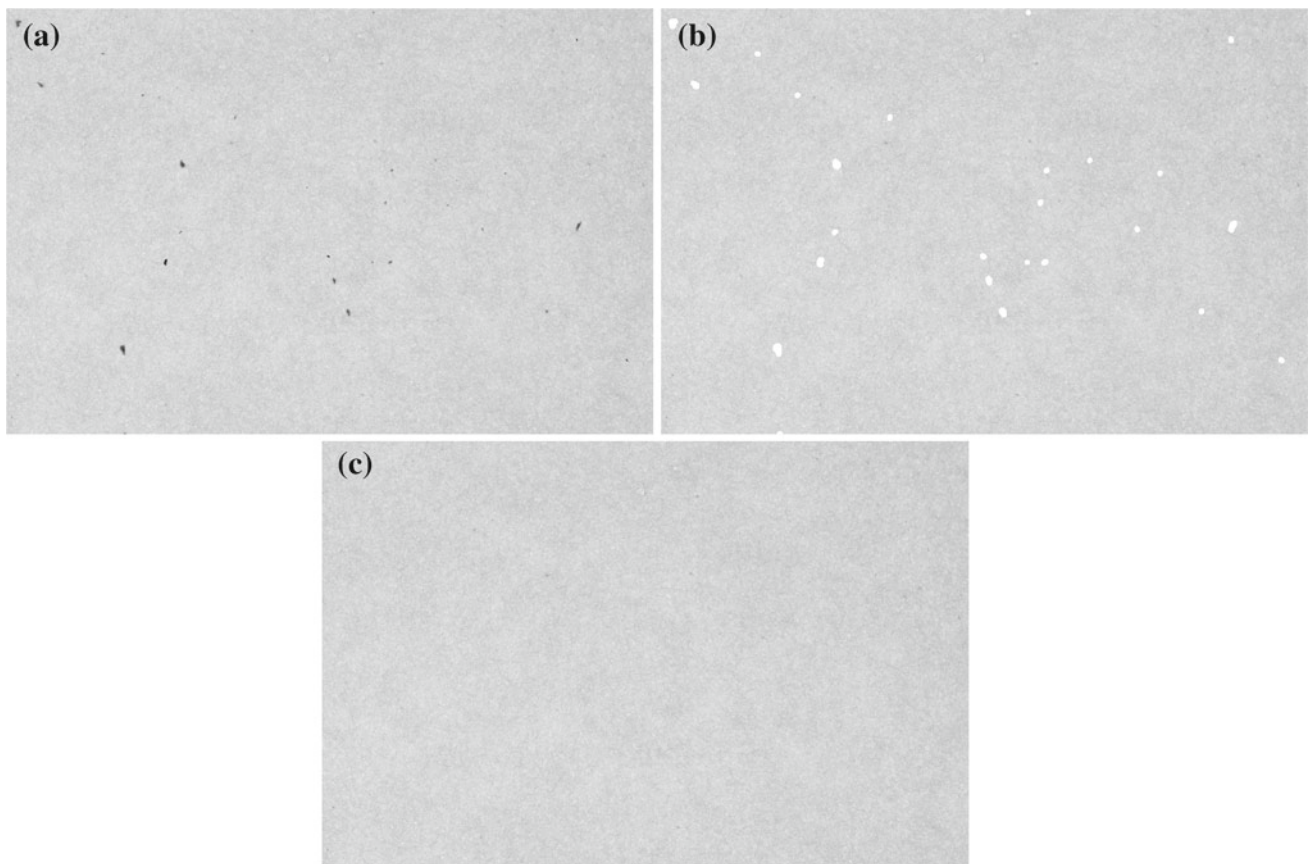


Fig. 4 Background generation: **a** initial target image, **b** target image with removed dirt segments, **c** semisynthetic background image

the reason for the different dirt counts is that in some cases the dirt particles are placed too close to each other and they are detected as a single particle. Since the background contains the synthetic parts, it can be a reason for losing some of the particles. One can also notice that for the larger particles, e.g., shives or sand, the difference in the counts is lower.

In addition, the experiments were performed to compare the total area of the segmented particles in pixels. Table 5 presents the results. One can see that not only are the counts of dirt different in the semisynthetic images, but also the area of the image segments as well as the geometric features. This means that either the features of some particles are modified, some particles can be lost or the wrong particles may be included in the class. This also shows that the greater losses occur in the cases of smaller particles, such as plastic and bark.

5.3 Dirt classification and the effect of the semisynthetic procedure on the classification

The experiments were performed on the softwood pulp after the second chlorine dioxide bleaching stage D_1 pulp. The

amount of the dirt particles per each class is presented in Table 6.

The results of the classification are presented in Fig. 6, including the performance of the classification methods on the training and test sets with separate results for different classes. For the 1NN classifier, the classification on the training set is trivial and, therefore, it is not shown. Figure 7 demonstrates how the data were modeled by the GMM classifier with the Figueiredo-Jain criterion.

As expected, the classification methods, which model the data, give results better than k-NN classification. For example, it can be seen from Fig. 8 that the boundaries between the classes are non-linear, whereas the LDA method models them as strict lines as shown in Fig. 9. On the other hand, SVM with the radial basis function kernel outperforms all other classification methods in almost all cases.

Figure 6 presents the classification results on the dirt particles extracted from the original images. The same features were employed as in the experiment with the semisynthetic data. It can be seen that using the semisynthetic procedure impairs the performance, yet allowing to obtain satisfactory results.

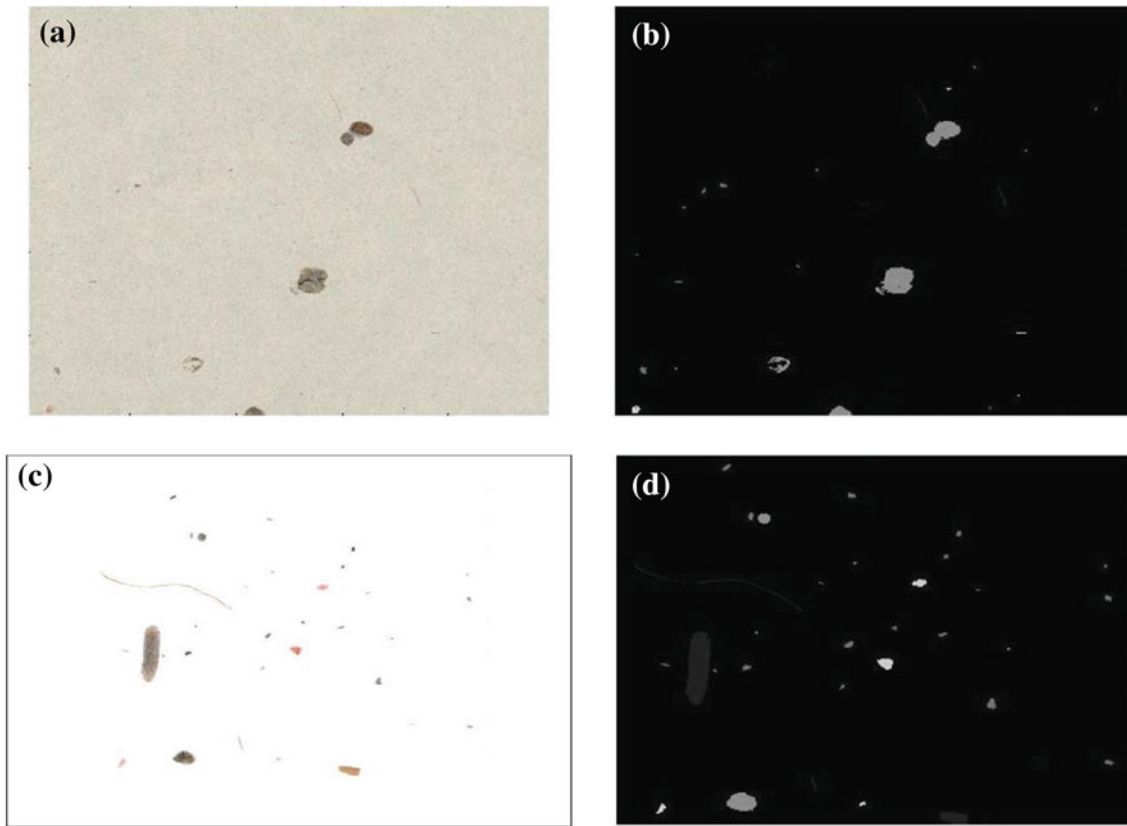


Fig. 5 Example images: **a** semisynthetic image with the unbleached pulp background, **b** ground truth mask for the unbleached pulp image, **c** semisynthetic images with the bleached pulp background, **d** ground truth masks for the bleached pulp image

Table 4 Dirt counts for original and generated images

Image	Bark	Plastic	Shive	Sand
Original	445	387	309	213
Semisynthetic	421	377	310	219

Table 5 Average segmented dirt area and average area difference for original and generated images

Image	Bark	Plastic	Shive	Sand
Original	83,830	89,312	127,611	384,700
Semisynthetic	75,293	75,515	117,324	376,780
Difference (%)	10.2 %	15.5 %	6.2 %	2.7 %

Table 6 Amount of particles for each dirt type

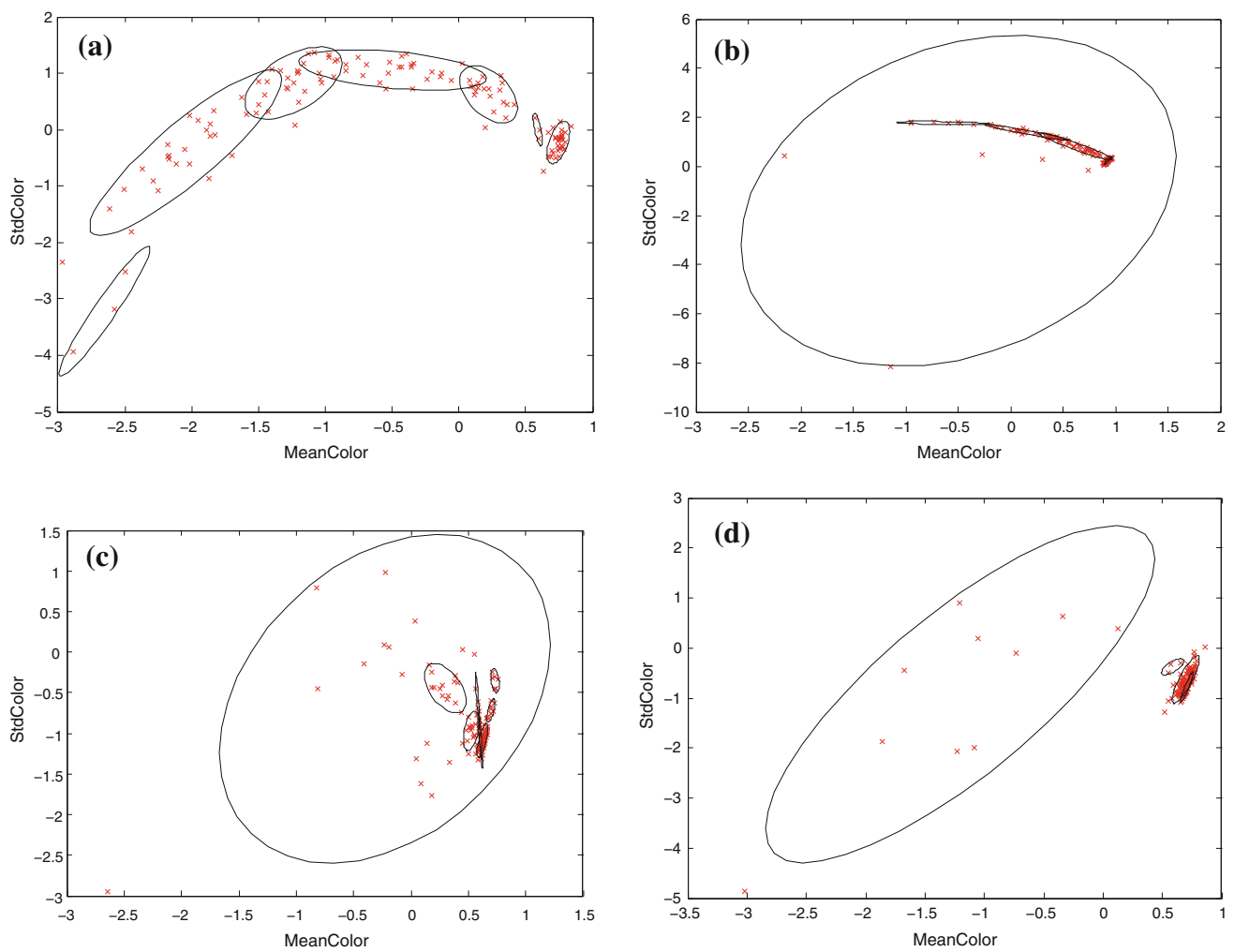
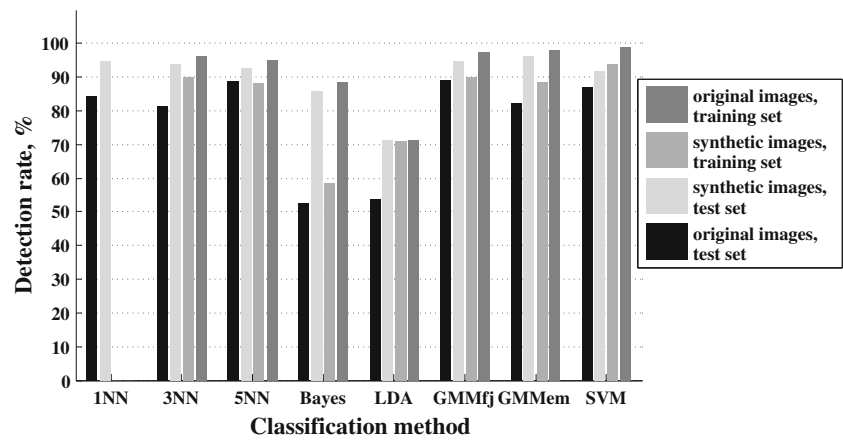
	Bark	Plastic	Sand	Shive
Training set	100	100	100	100
Test set	751	158	138	66

For the sake of visualization, the size of the feature set was restricted to two features. In this case, the set of features includes “MeanColor” and “StdColor”, which are illustrated in Fig. 8. The features were selected using the sequential search with a linear evaluation function.

5.4 Method performance when an unknown dirt type appears

Feature evaluation plays an important role when new type of dirt appears in the production. The close-to-optimal feature set, used in the classification, should be updated according to the dirt type that appears. To demonstrate how the classification results change if the feature set is not updated each time the new dirt type appears, the experiments were performed according to Algorithm 4.

For each set of dirt types, the classification results of the close-to-optimal feature set and previous close-to-optimal feature set are compared. The amounts of dirt particles in the training and test set are presented in Table 8. The results of feature evaluation and selection can be found in Table 7,

Fig. 6 Classification results**Fig. 7** Classes modeled with the GMM classifier with the Figueiredo-Jain criterion: **a** bark, **b** plastic, **c** sand, **d** shive

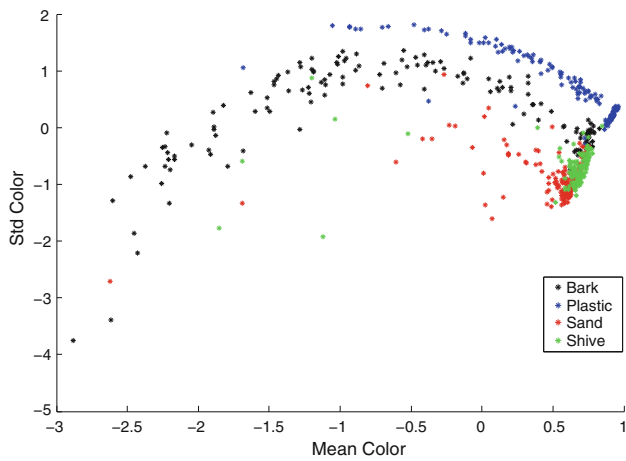


Fig. 8 A close-to-optimal feature set

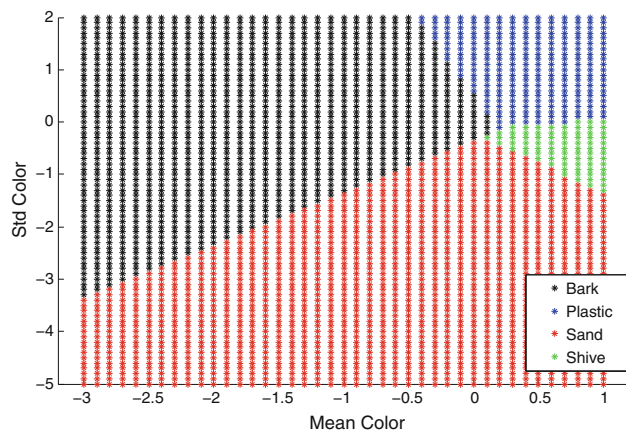


Fig. 9 Data boundaries representation by the LDA analysis

Algorithm 4 Feature selection experiment procedure

1. Select the close-to-optimal feature set for the classification of two types of dirt.
2. Repeat the following steps until all the types of dirt participate in the classification:
 - 2a. Add another type of dirt.
 - 2b. Perform the classification of the new set of dirt types using the previously selected features.
 - 2c. Select the close-to-optimal feature set for the new set of dirt.
 - 2d. Perform the classification of the new set of dirt types using the new feature set.
3. Compare the results of the classification using the previous and the new close-to-optimal feature sets.

which shows how the close-to-optimal feature set changes after the new dirt appears or disappears.

Figure 10 presents the detection rates in the situation when the new class appears. The percentage of correctly classified dirt particles is higher, when the close-to-optimal feature set is used than when the feature sets are not updated.

Table 7 The selected features

Dirt types to be classified	Close-to-optimal feature set
Bark vs plastic	Coarseness
Bark vs sand	Coarseness
Bark vs shive	Extent, FormFactor, MeanColor
Plastic vs sand	ConvexArea, coarseness
Plastic vs shive	MaxDiameter, MeanColor, Std-Color
Sand vs shive	AspectRatio, curl
Bark vs plastic vs sand	MaxDiameter, MeanIntensity, coarseness
Plastic vs sand vs shive	MaxDiameter, MeanColor, Std-Color
Bark vs sand vs shive	Coarseness, MeanColor, StdColor
Bark vs plastic vs shive	Coarseness, MeanColor, StdColor
Bark vs Plastic vs shive vs sand	MaxDiameter, MeanColor, Std-Color, coarseness, AspectRatio

Table 8 Amount of particles for each dirt type

	Bark	Plastic	Sand	Shive
Training set	100	100	100	100
Test set	969	244	233	201

5.5 Segmentation and classification of dirt particles in the real pulp sheets based on the semisynthetic training data

The proposed approach was tested with real independent images of dried non-bleached pulp sheets with dirt particles, marked by an expert. “Real” in this context means that the images contain the dirt particles from the real process, but not manually selected or synthetically generated. The set consists of eight images. Examples of the images are presented in Fig. 11.

The expert marked only those dirt particles which he was fully confident to be a dirt particle. From the presented images it can be seen that there are other dirt particles in the sheets that were not annotated by the expert since the expert was less confident about them. The total number of the marked particles was 69 including 57 shives and 12 bark particles. The system was trained on the semisynthetic particles of bark and shives. Each training set of a class consisted of 150 particles. The set of features for classification consisted of “Mean-Color” and “Roundness”. The classification of the marked particles was performed with 82 % accuracy. Lower accuracy compared to the experiments with semisynthetic images can be explained by different imaging conditions causing different appearance of the dirt particles and mistakes made by the expert. It should be also noted that the segmentation method provides a larger amount of the detected particles. This happened since the expert could not decide about the class of each

Fig. 10 Classification of results when the new classes appear

	Detection rate in %			
	3NN	LDA	GMMem	
Bark vs Plastic Feature set: Coarseness close-to-optimal	100	90	100	training set
	98.4	94.8	97.6	test set
↓ Sand class added				
Bark vs Plastic vs Sand Feature set: Coarseness	95.3	84.6	91.3	training set
	80.7	76	63.4	test set
↓ Features reestimated				
Bark vs Plastic vs Sand Feature set: Coarseness, MeanIntensity, MaxDiameter close-to-optimal	84.6	53.3	100	training set
	57.6	45.5	81.5	test set
↓ Shive class added				
Bark vs Plastic vs Sand vs Shive Feature set: Coarseness, MeanIntensity, MaxDiameter	84.5	47.3	98.5	training set
	54.3	37.5	78.5	test set
↓ Features reestimated				
Bark vs Plastic vs Sand vs Shive Feature set: Coarseness, AspectRatio, MeanColor, MaxDiameter, StdColor close-to-optimal	98	58.3	100	training set
	94.2	52.2	95.2	test set

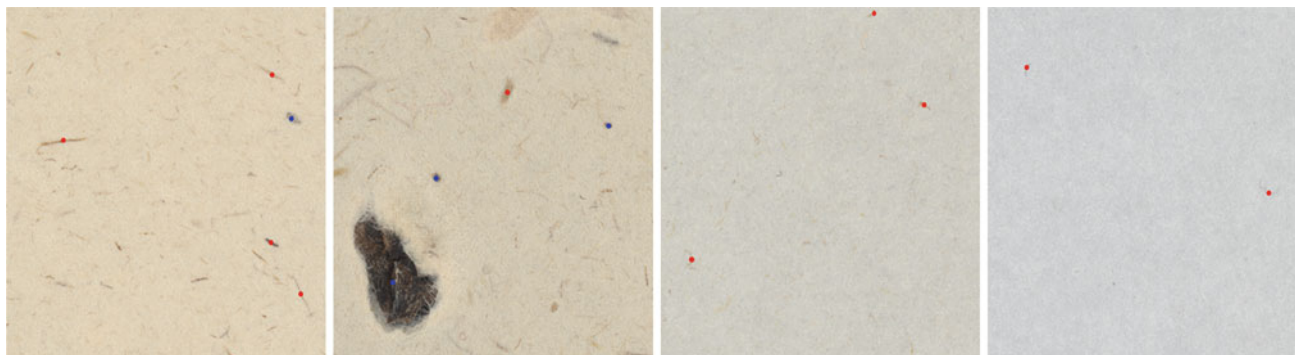


Fig. 11 Examples of the pulp sheets with real dirt particles. Bark particles are marked with *blue*. Shives are marked with *red*

single particle in an image and marked only those which he was fully confident.

6 Conclusion

This paper proposed a framework for developing dirt particle classification systems. The use of the framework begins with the problem of the ground truth generation and finishes with the analysis of the performance of the standard classification methods. Using the presented procedure for ground truth generation, there is no need for the manual annotation of the particles. The most important matter is that all of the particles that are contained in the semisynthetic images are

labeled and no unknown particles exist. The results proved that the semisynthetic procedure does not significantly affect the classification and segmentation results.

To make the system adaptable to the changes in dirt particle types, there is a feature evaluation stage where the most important features are determined. They represent the classes in the most efficient way. The experiments showed that the classification results improve significantly if the close-to-optimal feature set is used in the classification. The experimental part of the work presented the classification results for state-of-the-art and standard classifiers. It was shown that the methods modeling the data, such as GMM and SVM, outperformed other standard methods such as k-NN.

To justify our framework, experiments were performed also with unprocessed pulp sheets, i.e. dirt particles present in the sheets were from the real process, not manually added. With 82 % accuracy, the dirt particle classification method was shown to provide good results in a real industrial dirt identification task. However, the expert was able to annotate only a small portion of the dirt particles and was prone to mistakes. This motivates to use our framework for method development purposes. Using the framework it is possible to create reliably and comprehensive ground truth consisting real dirt particles even if the dirt types are visually indistinguishable to experts. Moreover, the amount of time-consuming annotation work can be minimized using the framework.

Acknowledgments The research was carried out in the “PulpVision” project (TEKES projects No. 70010/10 and 70040/11) funded by the European Union and the participating companies. The authors wish to acknowledge TEKES, and the companies for their support and collaboration.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Campoy, P., Canaval, J., Pena, D.: Inspulp: an on-line visual inspection system for the pulp industry. *Computers Ind.* **56**, 935–942 (2005)
- Dence, C.W., Reeve, D.W. (eds.): *Principles and Practice*. TAPPI Journal, Pulp Bleaching (1996)
- Drobchecko, A., Kamarainen, J.-K., Lensu, L., Vartiainen, J., Kälviäinen, H., Eerola, T.: Thresholding-based detection of fine and sparse details. *Front. Electr. Electron. Eng. China* **6**(2), 328–338 (2011). doi:10.1007/s11460-011-0139-x
- Duda, R., Hart, P.: *Pattern Classification*. Stork. Wiley, New Jersey (2001)
- Fastenau, H., Hagedorn, J., Jousimaa, T.: Automatic dirt counting during the production of market pulp. *TAPPI J* **74**(6), 73–78 (1991)
- Figueiredo, M., Jain, A.: Unsupervised learning of finite mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(3), 381–396 (2002)
- Fouladgaran, M., Mankki, A., Lensu, L., Käyhkö, J., Kälviäinen, H.: Automated counting and characterization of dirt particles in pulp. In: *Proceedings of the International Conference on Computer Vision and Graphics (ICCVG)*, Warsaw, Poland, **6374**, 166–174 (2010)
- Fukunaga, K.: *Introduction to Statistical Pattern Recognition*. Academic Press, London (1990)
- Gullichsen, J., Levlin, J.: *Chemical Pulping*. Papermaking Science and technology, Fapet Oy (1999)
- Juntunen, P., Tornberg, J., Ailisto, H.: Automated analysis of coloured ink particles in recycled pulp by machine vision. *Paperi ja Puu* **81**(5), 375–378 (1999)
- Kittler, J., Illingworth, J.: On threshold selection using clustering criteria. *IEEE Trans. System Man Cybern. SMC* **15**, 652–655 (1985)
- Komodakis, N., Tziritas, G.: Image completion using efficient belief propagation via priority scheduling and dynamic pruning. *IEEE Trans. Image Process.* **16**(11), 2649–2661 (2007)
- Levlin, J.-E., Söderhjelm, L.: *Pulp and Paper Testing*. Papermaking Science and technology. Fapet Oy, Helsinki (1999)
- Molina, L., Belanche, L., and Nebot, A.: Feature selection algorithms: A survey and experimental evaluation. In: *Proceedings of the 2002 IEEE International Conference on Data Mining* (2002)
- Parker, S., Chan, J.R.: Dirt counting in pulp: An approach using image analysis methods. In: *Proceedings of the IASTED International Conference on Signal and Image Processing (SIP)*, Kauai, Hawaii, USA (2002)
- Pulps - estimation of dirt and shives - part 1: Inspection of laboratory sheets by transmitted light. ISO 5350-1 (2006)
- Pulps - preparation of laboratory sheets for physical testing - part 1: Conventional sheet-former method. ISO 5269-1 (2005)
- Rosenberger, R.: Using a spread sheet model to develop and evaluate thresholding setting techniques for digitized image dirt counting systems. *Prog. Paper Recycl.* **4**(3), 39–54 (1995)
- Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. *J. Electron. Imaging* **13**(1), 146–165 (2004)
- Sinha, A., Gupta, S.: A fast nonparametric noncausal MRF-based texture synthesis scheme using anovel FKDE algorithm. *IEEE Trans. Image Process.* **19**(3), 561–572 (2010)
- Strokina, N., Eerola, T., Lensu, L., and Kälviäinen, H. Adaptive classification of dirt particles in papermaking process. In: *Proceedings of the 17th Scandinavian Conference on Image Analysis (SCIA2011)*, Ystad, Sweden (2011)
- Strokina, N., Mankki, A., Eerola, T., Lensu, L., Käyhkö, J., Kälviäinen, H. Semisynthetic ground truth for dirt particle counting and classification methods. In: *Proceedings of the 12th IAPR Conference on Machine Vision Applications (MVA2011)*, Nara, Japan (2011)
- Sutman, F.J.: Sampling statistics applied to automated tappi dirt counting. *TAPPI J.* **77**(5), 179–182 (1994)
- Theodoridis, S., Koutroumbas, K.: *Pattern Recognition*. Academic Press, London (1999)
- Verity IA eBusiness Team: The VERITY IA color image analysis software. <http://www.verityia.com/dirt-counter.php> (2007). Accessed 15 Jan 2011
- Wei, L. Y., Levoy, M.: Fast texture synthesis using tree-structured vector quantization. pp. 479–488
- Zeyer, C., Venditti, R., Puangchinda, K.: The distribution of impurities in pulp and paper—the effects of the random distribution of impurities on image analysis. *TAPPI J.* **78**, 168–175 (1995)
- Zhang, J., Fieguth, P., Wang, D.: Random field models. *Handbook of Image and Video Processing*, pp. 301–312 (2000)