2002 BELTWIDE COTTON CONFERENCES, ATLANTA, GA – JANUARY 8-12

HIGH SPEED TRASH MEASUREMENTS

M. Siddaiah, M.A. Lieberman and S.E. Hughs
Southwestern Cotton Ginning Research Laboratory
USDA-ARS
Mesilla Park, NM

Abstract

This paper discusses the identification of trash objects in cotton using machine vision-based systems. Soft computing techniques such as neural networks and fuzzy inference systems can classify trash objects into individual categories such as bark, stick, leaf, and pepper trash types with great accuracies. High speed trash measurements, enables the implementation of these techniques for on-line identification of trash. This identification of trash objects to individual categories can be used for the dynamic allocation of trash extraction equipment during the ginning process. Such a system can be implemented in a modern gin, to configure an optimal set of equipment during ginning to produce quality cotton. Classification of cotton in real-time allows for an automated means for assignment of trash grades to cotton, and could have a significant impact on the cotton industry.

Introduction

The cotton industry in the United States is developing new techniques to produce quality cotton and stay competitive in the world market. One of the objectives of the Southwestern Cotton Ginning Research Laboratories (SWCGRL), Imaging Group, has been to develop an accurate and reliable trash measurement system that can be used by the industry. Such a system is currently under development to identify trash objects found in cotton. This system has the capability to process cotton images at high speeds for possible use in the on-line identification of trash. It also has the capability to measure trash content for predicting trash grades for ginned cotton. The system being developed, acquire images using a RGB color camera and the Matrox® family series imaging boards that feature on-board, high-speed processing. The Matrox Imaging Library (MIL Version 6.1) image processing software, along with algorithms developed at the SWCGRL are used to process cotton images to identify trash. The trash objects present in cotton can be identified as bark, stick, leaf, or pepper trash and the overall trash content in ginned cotton can be computed.

Cotton Quality

Quality of cotton available for commercial use depends on a wide variety of factors (physical attributes). Variety, environment, harvesting procedures, and ginning practices affect these physical attributes. In order to maintain uniform cotton quality the United States Department of Agriculture-Agricultural Marketing Service (USDA-AMS), has established standardized procedures for measuring these physical attributes. These attributes are used to assign class grades to cotton as an aid to marketing and final commercial use. Classification of cotton is based on fiber length, length uniformity, fiber strength, micronaire, color, and trash (USDA-AMS, 1999). The presence of trash and its subsequent removal during the ginning process affects some fiber properties, trash grade, and color grade (Rogers, 1984). The presence of trash in ginned lint also impacts the spinning and textile manufacturing industries. As such, trash extraction is a vitally important aspect in the production of quality cotton.

Trash in Cotton

All commercially grown cotton in the United States is harvested mechanically. Two types of harvesting equipment are used to harvest the cotton crop, the spindle picker (machine picked), and the cotton stripper (machine stripped). In spite of advances in harvesting equipment, large quantities of trash are collected along with the cotton. This trash for the most part consists of plant foliage. The amount of trash is significantly higher for stripper-harvested cotton compared to picker-harvested cotton.

The primary objective of a commercial gin is to separate cotton fiber from the cottonseeds and trash. When cotton is harvested even under ideal conditions, pieces of leaf and other trash are picked along with the seedcotton. The cotton is passed through the requisite number of seedcotton cleaning and extraction machinery to extract large trash objects. The gin stand then separates the lint from the seeds. Material left behind in the lint can consist of bark, stick, leaf, pepper trash, grass, hulls, seed coat fragments, and motes. As the name suggests, bark is the outer covering of sticks and is rather stringy in appearance and not easy to separate from the fibrous lint material. Pepper trash refers to small broken or crushed pieces of leaf and other trash objects. Hulls are the outer coverings of the cotton boll. Immature cottonseeds are referred to as motes. The lint is then further cleaned before being baled. In spite of advances in cleaning and ginning equipment, a large number of trash objects remain in the ginned cotton.

Trash Identification

The extraction of shape descriptors or features that distinguish the trash objects is an important aspect in the identification of trash types, since features extracted from cotton samples form the feature set (patterns) for various classifier design. Classing specialists when assigning grades can visually inspect cotton samples and assess the various trash objects present in ginned cotton and, determine the overall distribution of the trash types. They can categorize the trash types as bark, stick, leaf, and pepper trash based on various physical attributes or properties inherent to individual trash types. For example, a bark object can be easily distinguished from other trash types by its stringy appearance. Stick objects are characterized by their oblong appearance (narrow and long in nature). Leaf objects can be distinguished from pepper trash based on the size (area) of the object. However, to categorize the trash types using machine-vision systems, unique identifiers are needed to distinguishing the trash types. These unique identifiers extracted for each trash type can result in higher classification rates and improve the overall performance of the classifier.

Figure 1 is a schematic of the various steps involved in the identification of trash. The first step in this process is to acquire raw cotton sample images. The acquired images are flat field corrected for any spatial illumination non-uniformity. The next step is the segmentation of cotton images to obtain a thresholded image. These images are binary images where trash pixels are identified and separated from lint background. The threshold level and the segmentation technique used to separate trash pixels from lint background, dictates the classification accuracy's of the trash types. Poor segmentation can separate a single piece of trash object into many artifacts resulting in misidentification of the trash types. This in turn will lead to poor classification rates. For example, a large piece of bark may have its whiskers (objects that are striped from the outer covering of stick objects) segmented as stick or pepper trash types. In certain instances, the object may be partially covered by lint resulting in discontinuities in the segmented image. If the threshold level is higher, trash objects that are close to each other can be segmented as a single object in the thresholded image.

The next step involves the transformation of the segmented binary image into a border kill image. In order to identify trash types accurately, objects lying on the boundary of the image are removed from the binary image. This process eliminates trash objects lying on the boundary of an image from being misclassified. For instance, a stick object with only a small area in the field of view of the camera can be misclassified as leaf or pepper object.

The border kill image is then converted to a label image. In a label image, the various trash objects in the image are assigned an ordinal number. The pixel values of the object are changed to the ordinal number giving different color for each object in the image when displayed. These objects are generally referred to as blobs. The label image is used by the image analysis software along with the raw cotton image (unsigned char image) to provide the various parameters for each blob in the label image. The software provides a means to include or exclude blobs based on the minimum and maximum limits of the various feature measures resulting in the parameters collected for the objects of interest. All blobs with area less than 10 pixels (1270 microns equivalent diameter) are filtered from the label image. These objects are considered as noise.

Image Acquisition

Most image processing applications involve acquisition of images for processing and extraction of information based on area of interest. In most applications, the imaging surface is a homogenous surface (flat or rigid). However, acquisition of cotton images poses certain problems that are inherent to cotton due to its physical properties. Since cotton is fluffy and fibrous in nature, cotton image when acquired in its free form, contain holes and perforations. Holes if present in the image can cause segmentation defects when trying to separate trash pixels from lint background. These holes will appear as trash objects in the segmented image.

For acquisition of cotton images, it is customary to press a known volume (85 grams) of cotton against a glass plate and the camera mounted across the glass plate. This ensures that the acquired image has a flat or an even (homogenous) surface, with the lint and the trash objects captured as they appear on the surface of the cotton samples. This is generally achieved by increasing the density of the free form cotton with pneumatically driven compression ram that presses cotton against the glass plate. It is observed in general practice that the images acquired are free of holes or perforations.

The cotton sample surface is illuminated by a light source for acquiring consistent, good quality, color image (RGB) for processing. It should be noted that the use of a light source to illuminate the imaging surface poses reflectance and aberration effects with glass plate covering the surface of the cotton sample. Based on previous experiences at SWCGRL, use of non-reflective glass plates has shown that the acquired images are free of reflection and aberration effects.

Figure 2 shows the setup that is a modified design of the trashmeter developed by Motion Control Inc® (MCI). These MCI machine was previously used by AMS, for both trash and color measurements. In this setup, a regulated power supply provides illumination via quad fiber-optic bundles with focusing lens at the light exits. The power source used is a DC regulated, light-feedback fiber-optic light source manufactured by Illumination Technologies®. It provides a highly regulated, cool white light at high intensities for injection into the fiber-optic light guides. It uses a 150-watt tungsten halogen lamp. The surface is illuminated by placing the quad lights opposite to each other at an angle of 45 degrees from the imaging plane. The camera used is a SONY® with pedestal adjustment for both black and white calibration adjustments.

Flat-Field Correction

Flat-field correction (Lieberman, and Patil, 1997) is a process wherein the intensity of a pixel located in an area of lower illumination is increased and the intensity of a pixel in an area of higher illumination is decreased. The intensity level of each pixel is multiplied by the ratio of the average intensity of the reference tile image (white tile) to the intensity of the respective reference pixel. If $I_{x,\,y}$, represent the intensity level of a pixel at the location x, y, then the intensity of a flat-field corrected image is represented by

$$I_{x,y}^{(\text{out})} = I_{x,y}^{(\text{in})} \frac{I^{(\text{tile\_avg})}}{I_{x,y}^{(\text{tile})}}$$

where

$I_{x,y}$ = intensity at position x, y

out = output image pixels

in = input image pixels

tile = reference image pixels

tile_avg = average gray level for the reference tile.

Training Sample Preparation

Training samples for measuring the various features are prepared by placing trash objects on bleached cotton. Each sample contains a single type of trash namely, bark, stick, leaf, or pepper trash. Two sets of training samples were prepared for each type of trash. Figures 3 through 6 illustrate the training samples for bark, stick, leaf, and pepper trash types and the segmented images. Trash objects were chosen from a bale of cotton to include all possible extremes with regards to size and shape of the trash types. The distribution of trash types in the training samples is not relative to their distribution in the entire cotton bale. For example, the amounts of pepper trash present in cotton bales are significantly higher compared to the presence of bark or stick objects. In addition, depending on the quality of harvested cotton the trash types may differ in their shape and size. The illustration is a typical set of trash types for a given bale of cotton. The objects were chosen to get a good representation of all trash types found in ginned cotton. Trash types belonging to each category is removed from the cotton bale using tweezers. These trash objects are placed on prepared bleached cotton background. The training samples are then placed in plastic bags to prevent samples from soiling due to excessive handling.

Segmentation

The RGB color image is transformed to the hue (H), intensity or luma (L), and saturation (S) component images. Entropy measure on the intensity plane is used as the threshold to segment the intensity plane of the HLS image resulting in a binary image with trash pixels separated from lint background. The intensity plane image is the average of the red, green, and blue pixel values. If $I_{x,\,y}$ represents the pixel intensity at the location x, y in an image, then

$$I_{x,y} = \frac{R_{x,y} + G_{x,y} + B_{x,y}}{3} \qquad x = 1, 2, \cdots, 512, \quad y = 1, 2, \cdots, 480.$$

Feature Extraction

Feature extraction is one of the important aspects to trash identification. Extraction of unique identifiers for each type of trash can aid in the development of classifiers with higher classification rates. Numerous methods were evaluated to find these unique identifiers that form the feature space:

a) Features perceived by a classing specialist to be important to distinguish between the trash types.
b) Relevant default features provided by the imaging software.
c) Computed features that are suggested in literature, which could provide unique shape descriptors for individual trash objects (Russ, 1994, and Pratt, 1991).

For each object in the label image, a set of features was computed. These features form the basis for a component feature vector that defines a point in multi-dimensional feature space. The classifier partitions this multi-dimensional hyperspace into specific regions, where each region corresponds to a type of non-lint material or trash type. Each classification system forms a unique hypersurface separating the feature space into object regions. *Area*, *perimeter*, *shape factor*, *ferrets X_f and Y_f*, were the default features computed by the image analysis software. The features measured by the image analysis software, is for the entire collection of objects in the whole image. This is efficient in terms of the computation time required to collect the data.

Computed features:

Convex area: The *convex area* of the object is the area of a 32-sided irregular bounding polygon of the object. This area is also referred to as the *rubber band area*. This is due to the fact that, if a rubber band were to be fitted to the object it would engulf the contours of the object in the form of an irregular polygon. By locating the boundary pixels of the object, it is possible to identify the coordinates of the bounding polygon by using rotation equations (Russ, 1994). The convex area measured is used subsequently for the computation of *Solidity*.

Bounding box area: The *bounding box area* is the product of the ferrets X_f and Y_f. This is the area of a rectangle circumscribing the object with sides of the rectangle parallel to image edges. The bounding box area is used to compute the *Extent* measure of the object.

Solidity: Ratio of area to convex area.

Extent: Ratio of area to bounding box area.

There is no single feature that can identify any of the stated trash type except for pepper trash. All objects with area less than 200 pixels can be considered as pepper trash. *Area*, *perimeter*, *shape factor* and *aspect ratio* were previously used at the SWCGRL as features to identify the trash types (Lieberman and Patil, 1992, 1997). In order to obtain better accuracy and higher classification rates, various shape descriptors were analyzed for the different trash types. These shape descriptors, along with some of the default features form the feature set or patterns for the various classifiers. These features are used as inputs to the various classifiers to identify the stated trash types. *Area*, *perimeter*, and *shape factor* are the default features computed by the image analysis software. The shape descriptors measured for the various objects are *convexity*, *compactness*, *solidity* and *extent* and routines were developed by the SWCGRL to compute these measures.

Figure 7 illustrate the *convex perimeter*, *convex area*, and *bounding box area* measurements for typical bark, stick, and leaf trash objects commonly found in seed cotton. The objects shown are actual blobs processed by the image analysis software from cotton sample images. Table 1 shows the different trash types with *area*, *perimeter*, *convex perimeter*, *convex area*, *bounding box area*, *solidity*, and *extent* measures for typical bark, stick and leaf trash objects. Among all the features, bounding box area and extent measures are the only features that are rotation variant. Extent measure alone can provide a discriminatory feature to separate bark and stick objects from leaf and pepper objects. This is due to the variation of the bounding box area based on the orientation of the trash objects in the image plane. This variation is very small for leaf and pepper objects since these two trash types can be considered as circular objects.

For bark objects, there are two specific orientations anywhere between 0° and 90°, where the bounding box area is a minimum and a maximum. However, for stick objects the minimum is limited to two possible orientations. When a stick object is oriented either at an angle of 0° or 90°, the bounding box area is a minimum at this orientation. This area is much closer to the actual area of the stick object, resulting in extent values for stick objects closer to unity. The same object oriented at an angle of 45° has the maximum bounding box area resulting in low extent measure. The difference between extent measures at 0° and 45° (E^{dif}) provides a distinguishing feature for stick objects.

Bark objects are stringy in appearance. This is due to the stripping of the outer covering of the stems of cotton plant during harvesting and during ginning. Due to its stringy appearance, the convex perimeters of these objects are different from the actual perimeter of the objects. This results in the convex area of bark objects to be much larger than the actual area of the object. Solidity values for bark objects are typically less than 0.5. In the case of stick, leaf, and pepper objects the convex perimeter is much closer to the actual perimeter of the object. This results in convex area of these objects being closer to the actual area of the trash objects. Hence solidity measures for stick, leaf, and pepper objects are closer to unity.

Soft Computing

Soft computing is an innovative approach that deals with intelligent systems that combine a wide array of techniques, resources, and knowledge in their design. Most real world applications currently solved use a variety of techniques, which generally are exclusive of each technique. With the approach of soft computing, hybrid systems can be developed which use these techniques synergistically rather than exclusively, where each technique can complement one another in the development of these intelligent systems. This results in a developed system that is adaptive, more responsive to changing environments, and possess human like expertise within a problem domain to make decisions or provide reasoning for the decisions being made. Such systems are referred to as neuro-fuzzy systems. Neural networks that recognize patterns and adapt themselves to ever changing environments; fuzzy inference systems that incorporate human knowledge and make decisions based on some reasoning and inference are some examples of neuro-fuzzy systems.

Several computing paradigms, including neural networks, fuzzy set theory, approximate reasoning, and derivative free optimization methods like genetic algorithms fall in the category of soft computing. Each of these methodologies has their strengths and weaknesses, and the integration of these methodologies for any given problem domain is the basic essence of soft computing.

Results from three classifiers developed to identify the trash types commonly found in cotton are presented in this paper. The classifiers developed use artificial neural networks, fuzzy-logic based approaches, and hybrid systems to identify trash types in cotton samples. The three classifiers are:

a) Fuzzy clustering technique.
b) Back propagation neural network.
c) Adaptive Network-based Fuzzy Inference System also known as ANFIS.

The classifiers developed are tested for their effectiveness in identifying the trash objects to their respective categories. The classification results serve as a basis for use in the prediction of the trash content and categorization of the individual trash objects in cotton. This enables the ginner to measure the quantity of the individual trash types present in harvested cotton. The technique with the best classification results can be used to identify trash types in real-time (on-line) during the ginning process. Based on the trash types and the amount of each trash type, the ginning process can be optimized by selecting the desired level of drying and cleaning machinery during processing to produce optimal quality lint. This will eventually lead to the development of gin plants that are fully automated.

Trash objects in five test samples were classified to evaluate the performance of the three classifiers. Figures 8 and 9 illustrate two of the test samples. The classification results of the trash objects in the five test images are summarized in Table 2 (Siddaiah, et al., 1999abc). Results indicate that neural networks and ANFIS identify the trash objects with great accuracies. The misclassification of some of the objects is mainly due to defects during segmentation. Better segmentation of the cotton images can result in higher classification accuracy's.

Trash Content Measurement

Trash Content is defined as the ratio of the total trash area to the image area, i.e.,

$$\text{Trash Content } (\%) = \frac{\text{Trash Area}}{\text{Image Area}} \times 100.$$

The trash area is measured using trash meters. Trash meters count the number of pixels identified as trash based on a certain threshold criteria for a known area of the cotton sample. This measure is currently indicated on cotton classification sheets in addition to classer assigned trash grades during cotton classification. AMS would eventually like to replace the classer assigned trash grade once a reliable system is developed to predict trash grades for cotton. As such, the development of such a system is significant to the industry, since class grades could be assigned to cotton bales at the gin.

The trash content measures for 18 AMS box samples with varying trash levels were computed. These trash boxes are used as standards for calibration of trash meters. Figures 10 and 11 represent two such samples with the lowest and highest trash content.

Table 3 illustrates the AMS and SWCGRL trash content measures for the 18 box samples. It is seen that the SWCGRL measures were higher but consistent with the AMS measures. If an industry standard is developed to tie in the trash content measurements with the classer trash grades, the trash content measurements obtained by the imaging system developed at the SWCGRL can be used in the on-line assignment of trash grades for ginned cotton (Siddaiah, et al., 2000).

Conclusion

The paper presents a general framework for high speed trash measurements in cotton using a machine-vision based system. Soft computing techniques discussed indicate that trash objects commonly found in cotton can be categorized as bark, stick, leaf and pepper trash with great accuracies. The on-line identification of trash objects to individual categories could be used to configure an optimal set of cotton cleaning and extraction equipment. The implementation in a modern gin could result in the production of higher quality lint and subsequent finished consumer products. The computation of the overall trash content and its possible use in the assignment of trash grades to ginned cotton is economically attractive to the industry. The impact of an automated system is wide ranging and could have significant benefits to the entire cotton industry.

Disclaimer

Mention of a trade name, proprietary product, or specific equipment does not constitute a guarantee or warranty by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products that may be suitable.

References

United States Department of Agriculture, Agricultural Marketing Service. (Revised January 1999). The Classification of Cotton, Agricultural Handbook 566, Washington, D.C.

Lieberman, M.A. and R.B. Patil. 1992. Non-lint material identification using computer vision and pattern recognition. SPIE, Optics in Agriculture and Forestry, 1836:142-152.

Lieberman, M.A. and R.B. Patil. 1997. Evaluation of learning vector quantization to classify cotton trash. Optical Engineering, 36-3:914-921.

Pratt, W.K. 1991. Digital Image Processing. 2nd edition, John Wiley and Sons, New York.

Rogers, C.D. 1984. Trash meter response to trash removal in processing. Proceedings of the Beltwide Cotton Conferences, 361-364.

Russ, J.C. 1994. The Image Processing Handbook. 2nd edition, CRC Press, Boca Raton.

Siddaiah, M., M.A. Lieberman, S.E. Hughs, and N.R. Prasad. 1999. A soft computing approach to classification of trash in ginned cotton. Proceedings of the Eighth International Fuzzy Systems Association World Congress, 1:151-155.

Siddaiah, M., M.A. Lieberman, S.E. Hughs, and N.R. Prasad. 1999. Identification of trash types in ginned cotton using neuro fuzzy techniques. 1999 IEEE International Fuzzy Systems Conference Proceedings, FUZZ-IEEE'99, II:738-743.

Siddaiah, M., M.A. Lieberman, S.E. Hughs, and N.R. Prasad. 1999. Computation of trash content in ginned cotton using soft computing techniques. Proceedings of the 42nd Midwest Symposium on Circuits and Systems, 1:547-550.

Siddaiah, M., M.A. Lieberman, S.E. Hughs, and N.R. Prasad. 2000. Identification of trash types and correlation between AMS and SWCGRL trash content in ginned cotton. Proceedings of the 2000 Beltwide Cotton Conferences, 2:1549-1555.

Table 1. Feature measurements.

Trash type	Bark	Stick	Leaf
Area	1526	1578	581
Convex area	3123	1638	624
Bounding box area at 0°	4640	2079	900
Bounding box area at 45°	8692	5700	840
Perimeter	392.9700	238.5920	97.8000
Convex perimeter	274.7300	222.4640	95.1100
Solidity	0.4886	0.9634	0.9311
Extent at 0°	0.3289	0.7590	0.6456
Extent at 45°	0.1756	0.2768	0.6917
E^{dif}	0.1533	0.4822	0.0461

Table 2. Classification results.

Classifier	Total Number of Objects	Actual Trash Type	Classified Bark	Classified Stick	Classified Leaf	Classified Pepper	Classification Rate (%)
Fuzzy Clustering	10	Bark	0	4	6	0	88.57
	4	Stick	0	0	0	4	
	16	Leaf	0	0	2	14	
	215	Pepper	0	0	0	215	
Neural Network	10	Bark	6	4	0	0	90.20
	4	Stick	1	2	1	0	
	16	Leaf	0	3	6	7	
	215	Pepper	1	2	5	207	
ANFIS	10	Bark	6	4	0	0	95.10
	4	Stick	0	2	0	2	
	16	Leaf	0	2	14	0	
	215	Pepper	0	0	0	215	

Table 3. AMS and SWCGRL trash content measurements for the cotton trash boxes.

Sample No.	AMS Values	SWCGRL Values (avg. of 10 reps)	Ratio AMS to SWCGRL Values
# 1	0.1010	0.2191	0.4609
# 2	0.1066	0.2312	0.4611
# 3	0.2716	0.6258	0.4340
# 4	0.2727	0.5953	0.4581
# 5	0.2903	0.6316	0.4596
# 6	0.4209	1.0284	0.4093
# 7	0.4344	0.9574	0.4537
# 8	0.4422	0.9677	0.4570
# 9	0.4428	1.0374	0.4268
# 10	0.6791	1.4967	0.4537
# 11	0.7009	1.4068	0.4982
# 12	0.9006	2.1106	0.4267
# 13	1.1077	2.1989	0.5038
# 14	1.1573	2.3822	0.4858
# 15	1.4160	2.9832	0.4747
# 16	1.7370	3.4957	0.4969
# 17	1.906	3.9445	0.4832
# 18	2.8687	5.5055	0.5211

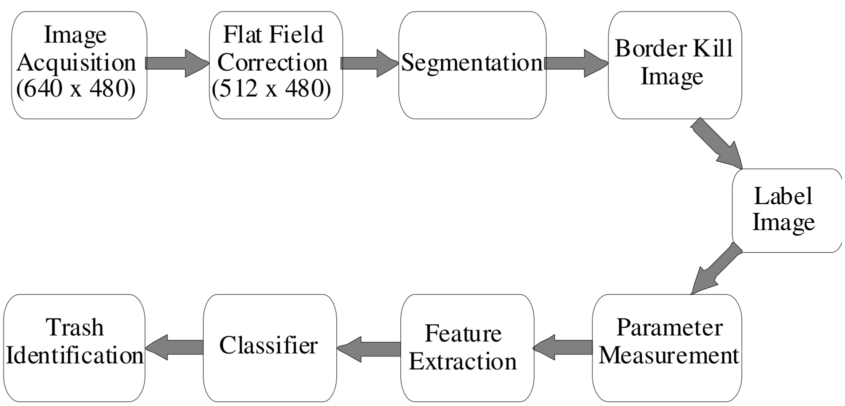
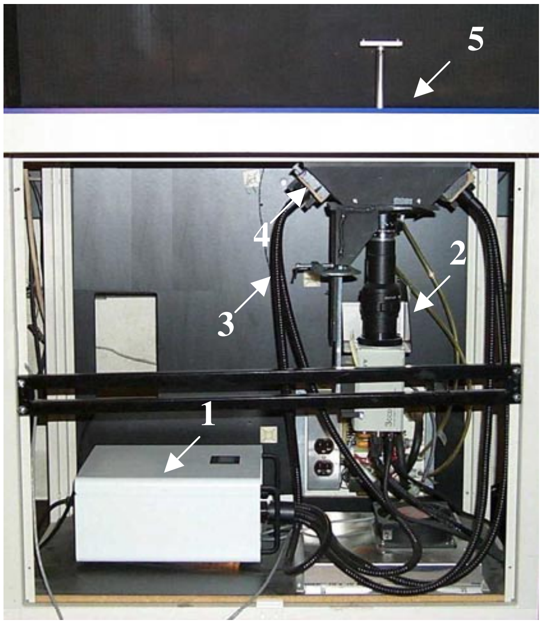


Figure 1. Steps involved in the extraction of features.



(1) Illumination Technologies light source; (2) RGB Sony color camera; (3) Quad fiber optic bundles; (4) Focusing lens; (5) Imaging window.

Figure 2. Setup to acquire cotton images using quad spot-lights with focusing lenses.

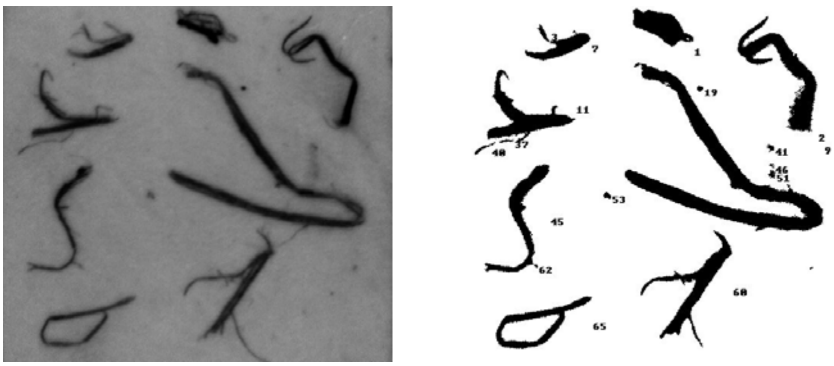


Figure 3. Bark training sample.

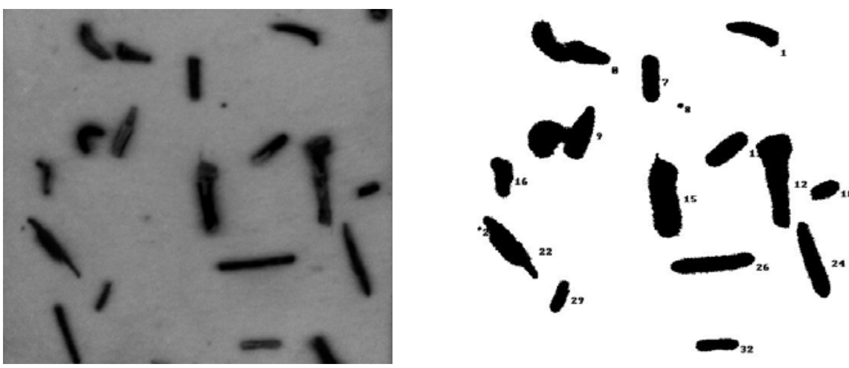


Figure 4. Stick training sample.

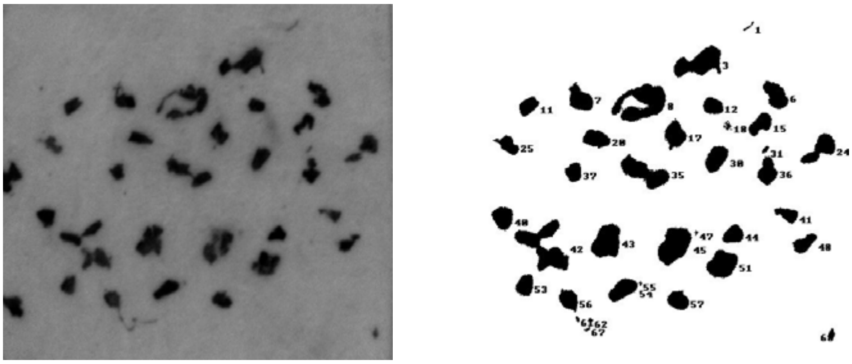


Figure 5. Leaf training sample.

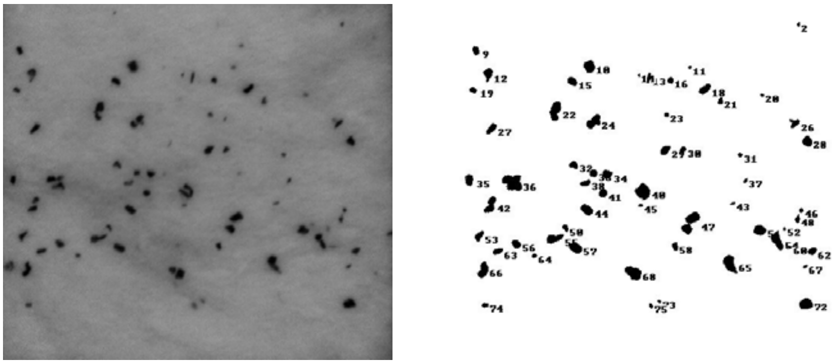


Figure 6. Pepper training sample.

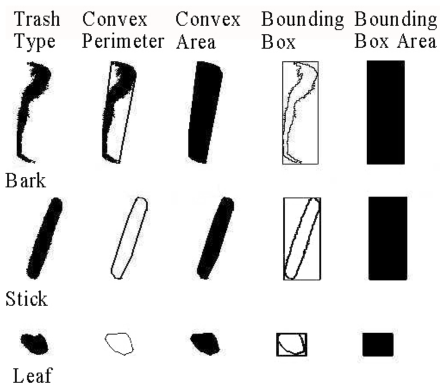


Figure 7. Trash types with feature measurements.

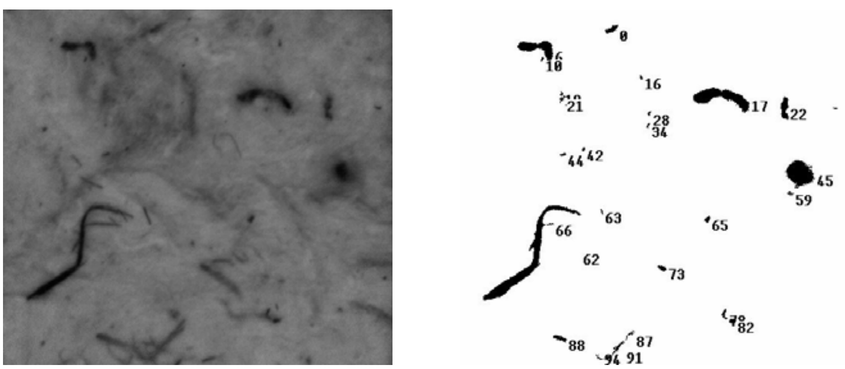


Figure 8. Cotton test sample #1.

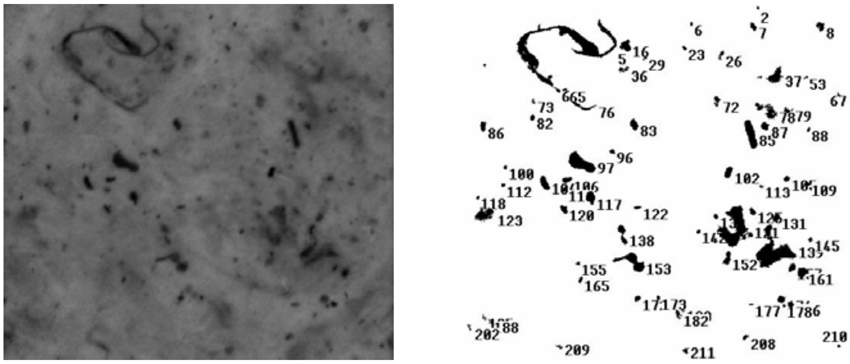


Figure 9. Cotton test sample #2.

Figure 10. AMS cotton box sample #1 (Trash Content = 0.27).

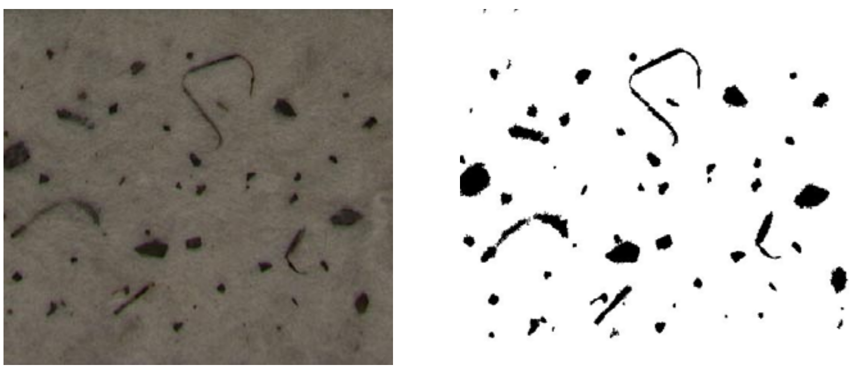


Figure 11. AMS cotton box sample #6 (Trash Content = 2.87).