Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Food Engineering 78 (2007) 597–605

JOURNAL OF FOOD ENGINEERING

www.elsevier.com/locate/jfoodeng

Stem and calyx recognition on 'Jonagold' apples by pattern recognition

D. Unay *, B. Gosselin

TCTS Labs., Faculté Polytechnique de Mons, Multitel Building, Avenue Copernic 1, Parc Initialis, B-7000 Mons, Belgium

Received 20 February 2005; accepted 28 October 2005
Available online 10 January 2006

Abstract

In this paper, a novel method to recognize stem or calyx regions of 'Jonagold' apples by pattern recognition is proposed. The method starts with background removal and object segmentation by thresholding. Statistical, textural and shape features are extracted from each segmented object and these features are introduced to several supervised classification algorithms. Linear discriminant, nearest neighbor, fuzzy nearest neighbor, support vector machines classifiers and adaboost are the ones tested. Relevant features are selected by floating forward feature selection algorithm. Support vector machines, which is found to be the best among all classification algorithms tested, correctly recognized 99% of the stems and 100% of the calyxes using selected feature subset. These results exhibit considerable improvement relative to the ones introduced in the literature.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Apple; Stem; Calyx; Machine vision; Quality sorting; Classification

1. Introduction

Tons of apples are produced, harvested and consumed throughout the world each year. Visual inspection of these apples are traditionally done by human experts. Even so, automatization of this process is necessary to increase speed of inspection as well as to eliminate human error and variation introduced by experts. In recent years machine vision systems have been widely applied to evaluate external quality of apples. However, these systems cannot provide robust and accurate results yet, because high variability of defect types and skin color as well as presence of stem/calyx (SC) areas increase complexity of the problem. Computer vision systems are mostly confused in discriminating SC ends from true defects due to their similarity in appearance. Hence, accuracy of apple sorting is diminished by false identification of SC ends.

Several approaches have been introduced to recognize SC's using mechanical or computer vision systems. Mechanical approaches include systems in which orientation of fruits, therefore positions of SC's are known. However, in reality, adjusting and preserving orientation of fruit reliably while acquiring images of whole apple surface is problematic. Moreover, in the image acquisition system used in this research as well as in most other systems introduced by other researchers, orientations of apples while imaging are not known. Hence, mechanical solutions are not considered in this paper.

Yang (1996) introduced an image analysis technique to identify SC's on 'Golden Delicious' and 'Granny Smith' apples. Assuming stems and calyxes appear as dark patches, first these areas are segmented by flooding algorithm from images under diffuse light. Then, 3D surfaces of patches are reconstructed from structural light projected image. Patches are classified as SC or patchlike defect by back-propagation neural network using features extracted from both images. Average recognition rate achieved was 95%, however, proposed method is tested only on monocolored apples. Crowe and Delwiche (1996a, 1996b) used structural illumination to detect apple defects, where concave dark spots were considered to be SC. Unfortunately, no numerical result was provided by the authors for

* Corresponding author. Tel.: +32 65 374745; fax: +32 65 374729.
E-mail address: unay@tcts.fpms.ac.be (D. Unay).
URL: http://www.tcts.fpms.ac.be (D. Unay).

0260-8774/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jfoodeng.2005.10.038

identification of SC's. Wen and Tao (1999) developed a rules-based NIR system and used histogram densities to discriminate SC's of 'Red Delicious' apples from defected areas. Recognition rates of stems and calyxes were 81.4% and 87.9%, respectively. Their system was less reliable when SC's appeared closer to the edge of fruit. In their later work, Wen and Tao (2000) used an NIR and a middle-infrared (MIR) camera for apple fruit inspection, where image from the latter was used to segment SC's and about 99% of them were correctly recognized. However, cost of cameras, which is not discussed by the authors, is an important issue for practical implementation of this approach. Penman (2001) illuminated four different varieties of apples with blue linear light and used reflection patterns of the fruit acquired by a ccd-camera to locate SC's as well as blemishes. Accuracy of the algorithm was inversely proportional with the location of SC's relative to fruit center. Moreover, the author mentioned about neither presence of defects nor their effect on recognition. Li, Wang, and Gu (2002) assumed that SC areas were concave and defected ones lost their concavity. So, fractal analysis with an artificial neural network is used to discriminate SC areas from defected ones in 'San Fuji' apples. Tests were done on a small database and it was reported that highly rotten areas were misclassified because their surfaces were concave. Cheng, Tao, Chen, and Luo (2003) proposed to use MIR camera with segmentation based on gray-level similarity of pixels to detect SC's on 'Red Delicious' apples. Euclidean distance was used to evaluate similarity. Recognition rates achieved for stems and calyxes were 94% and 92%, respectively. Kleynen, Leemans, and Destain (2005) utilized correlation-based pattern matching technique to detect SC's of 'Jonagold' apples in a multispectral vision system. Recognition rates for stems and calyxes were 91% and 92%, respectively. And 17% of defects were misclassified as SC's. Pattern matching method has been widely applied for object recognition, but its main disadvantage is its high dependency on the pattern (template) used. Recently, Bennedsen and Peterson (2004) used unsupervised feature extraction and neural networks to discriminate apple images including SC's from those that do not. Recognition rate achieved was 98%, however, their approach was not able to discriminate between true defects and stems or calyxes.

Above literature review reveals that SC identification is a necessary task for an accurate fruit sorting system, but it is not so easy to accomplish. Mechanical methods are simply not reliable enough. Computer vision-based methods introduced by other researchers cover wide range of materials and techniques, however, the quest for a general, accurate and cheap solution is still open. Pattern recognition on computer images is typically composed of object segmentation, features extraction, features selection and classification steps. And performance of overall system depends on individual accuracies of these subsequent steps. For example, a suitable feature selection process can improve performance by increasing recognition, removing irrelevant features or both. Similarly, certain classification algorithms can be more efficient in certain tasks than others. Hence, aim of this paper is to introduce a novel pattern recognition method for recognizing SC regions in 'Jonagold' apples images with special emphasis on feature selection step and effect of using different classification algorithms. 'Jonagold' variety is chosen due to their bi-colored skin, which increases the difficulty of the recognition problem.

2. Materials and methods

2.1. Image acquisition and database

Image acquisition device used for this research is simply composed of a high resolution (1280×1024 pixels) monochrome digital camera, four interference band-pass filters, a frame grabber, a diffusely illuminated tunnel with two different light sources (fluorescent tubes and incandescent spots), and a conveyor belt on which fruits are placed. The filters are centered at 450, 500, 750, and 800 nm with respective bandwidths of 80, 40, 80, and 50 nm. This device is capable of acquiring only one-view images of fruits. Each of these one-view images were composed of four filter images, which had to be separated by alignment based on pattern matching. Then, *flat field correction* is applied to remove vignetting on filter images. Finally, each filter image is composed of 430×560 pixels with 8 bits-per-pixel resolution (Fig. 1).

Database consists of images of 819 'Jonagold' variety apples, which are manually placed in the view of camera (Table 1). 280 of the images contain only healthy skin in view. 293 images are of stems or calyxes with various orientations with respect to the camera view. The rest of the images (246) have defects of various size and kind (russet, recent bruises of 1–2 h old, rot, scald, hail damage with and without perforation, scar tissue, limb rubs, ...). Bruises are produced by dropping the fruit from 30 cm height onto a steel plate. 'Jonagold' variety is selected, instead of mono-colored ones, because it has a bi-colored skin

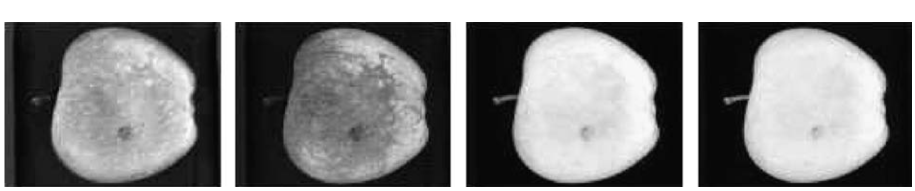

Fig. 1. Filter images of a fruit. Left to right: 450, 500, 750, and 800 nm filters.

Table 1
Image database

	No. of fruits
Healthy view only	280
Stem in view	148
Calyx in view	145
Defect in view	246
Whole database	819

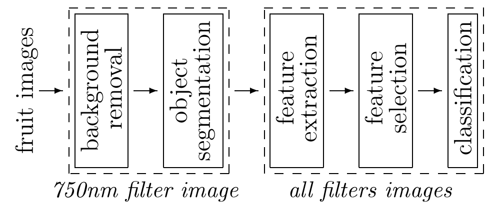

Fig. 3. Architecture of the proposed stem/calyx recognition system.

causing more difficulties in segmentation of objects due to color transition areas. Some images of the database can be observed in Fig. 2.

Construction of the image acquisition system and collection of the database were done by the Mechanics and Construction Department of Gembloux Agricultural University of Belgium. Therefore, for further details concerning image acquisition, filters selection and database, please refer to Kleynen, Leemans, and Destain (2003) and Kleynen et al. (2005).

2.2. Stem and calyx recognition

The proposed system to recognize SC of 'Jonagold' apples is composed of following steps: background removal, object segmentation, features extraction, features selection and classification as in Fig. 3. Before explaining these steps in the succeeding subsections, it has to be stated that in order to decrease computational expense of the whole process sizes of the images are first reduced to 128×128 pixels by nearest neighbor interpolation method.

2.2.1. Background removal

The database is composed of images of apple views on a dark, uniform colored (i.e., low intensity) background. Therefore, fruit area can be separated from background by thresholding the 750 nm filter image at intensity value of approximately 11.77%. Our visual observations have shown that fixed thresholding can remove low intensity regions like some defects, stems or calyxes. Hence, a morphological filling operation is applied to remove holes in fruit area caused by thresholding.

2.2.2. Object segmentation

Vignetting on the images, which is uneven illumination (transition from a brighter image center to darker corners) caused by image acquisition device, was removed with 'flat field correction' by Kleynen et al. (2005). However, our initial segmentation efforts revealed that segmentation was problematic at the far edges of fruit probably due to the still-existing illumination artifacts. Therefore, after background removal, fruit area is eroded by a rectangular structuring element of size adaptive to fruit size. Fig. 4 displays an example of this erosion process. Output binary image of erosion step is actually our region-of-interest (ROI) or region-of-inspection, in other words. Note that far edges of fruit area are removed from ROI and will not be inspected, which is not preferred because a visual inspection system has to examine full surface of the aimed object to be reliable. Although, current images of our database are from one-view, future goal of this research is to inspect full surface of fruits by multiple (4) views, which will hopefully overcome this limitation.

The ROI is then used as a mask to compute average (ρ) and standard deviation (ϵ) of intensity values of fruit. Subsequently, segmentation of objects (candidate SC's) is done by thresholding the masked fruit area with

$$T_0 = \rho - 2 \times \epsilon \qquad (1)$$

where pixels with intensity less than T_0 are believed to belong to an object. Finally, an adaptive spatial cleaning operation is applied to remove very small (smaller than 1% of ROI) objects and refine segmentation. Hence, the result is the binary segmentation image. This concludes segmentation of SC's.

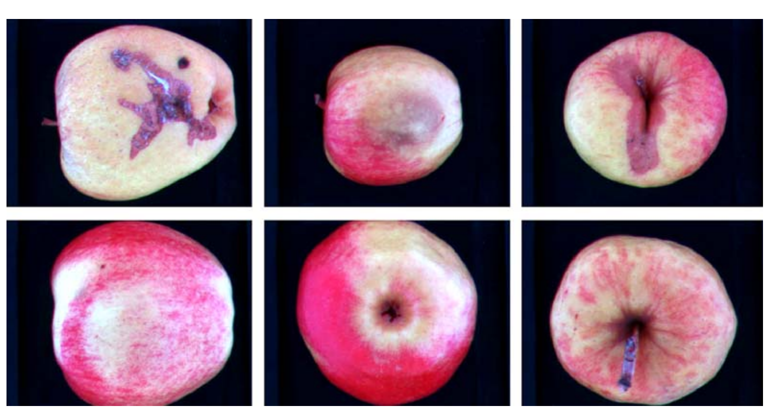

Fig. 2. Examples of original RGB images. Defected apples above and apples with healthy skin, calyx, stem, respectively below.

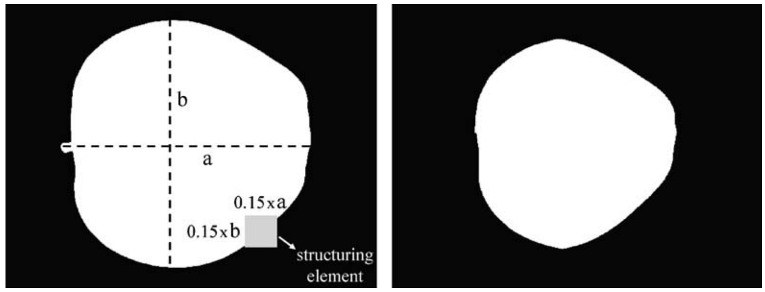

Fig. 4. Illustration of erosion process on the left and the corresponding result (ROI) on the right.

2.2.3. Features extraction

An ideal feature extractor should produce representations of objects to be classified as good as possible, and thus make the job of classifier as trivial as possible (Duda, Hart, & Stork, 2001). Moreover, long-term aim of this research is to provide a rapid algorithm for SC recognition, which means features to be extracted should be computationally cheap. Therefore, 7 statistical, 1 textural, and 3 shape features are extracted from each segmented object (Fig. 5). As statistical and textural ones depend on pixel intensity values, their computation is repeated with each filter image. In the end, each object is represented by a total of 35 features.

Performance of a classification algorithm will be biased if the features it uses are not scaled properly. Hence, the features are normalized to have a mean of 0 and standard deviation of 1.

$$
\textit{statistical}\begin{cases}
\text{average } (\mu) & = \dfrac{1}{N}\sum_{i=1}^{N} p_i \\
\text{standard deviation } (\sigma) & = \Big(\dfrac{1}{N-1}\sum_{i=1}^{N}(p_i-\mu)^2\Big)^{1/2} \\
\text{minimum } (min) & = min(p_i) \text{ for i=1,...,N} \\
\text{maximum } (max) & = max(p_i) \text{ for i=1,...,N} \\
\text{gradient } (grad) & = max - min \\
\text{skewness } (skew) & = \dfrac{\sum_{i=1}^{N}(p_i-\mu)^3}{N\sigma^3} \\
\text{kurtosis } (kurt) & = \dfrac{\sum_{i=1}^{N}(p_i-\mu)^4}{N\sigma^4}
\end{cases}
$$

$$
\textit{textural}\begin{cases}
1^{st} \text{ invariant moment } (\phi_1) & = \eta_{20}+\eta_{02} \text{ where } \eta_{xy} \text{ is the} \\
\text{of Hu (1962)} & \text{normalized central moment}
\end{cases}
$$

$$
\textit{shape}\begin{cases}
\text{area } (S) & = N \\
\text{perimeter } (P) & = N_p \\
\text{circularity } (C) & = \dfrac{P^2}{4\pi S}
\end{cases}
$$

Fig. 5. Details of features extracted. N is the number of pixels in segmented object. p_i refers to the intensity of ith pixel. N_p is the number of pixels in object perimeter. min(·) and max(·) refer to minimum and maximum of enclosed arguments, respectively. (See Hu (1962).)

2.2.4. Features selection

In real-world problems, relevant features are generally not known beforehand, which results in extraction of several features that also include irrelevant/redundant ones. Irrelevant or redundant features may introduce noise, thus have negative effect on accuracy of a classification algorithm. Furthermore, by using fewer features computation cost of the system can be significantly reduced. Therefore, feature selection is necessary to find a subset of best discriminating features by removing irrelevant/redundant ones.

Exhaustive search of feature space guarantees optimal solution, but for most of the real-world problems it is impractical as there exists 2^n possible sub-sets with n being the number of features. Jain and Zongker (1997) has divided statistical pattern recognition based feature selection algorithms into categories by being optimal/suboptimal, deterministic/stochastic and giving single/multiple solutions. They stated that suboptimal, single solution, deterministic methods (also referred as 'sequential' methods) are the most used ones for performing feature selection. Methods of this category begin with single solution and iteratively add or remove features until a termination criterion is met.

Sequential floating forward selection (SFFS) method of Pudil, Ferri, Novovicova, and Kittler (1994), which belongs to this category, was found to be a good choice for SC recognition (Unay & Gosselin, 2004). Therefore, SFFS method is used in this work. The algorithm starts with an empty feature subset. At each iteration, it tentatively adds one feature that is not already selected to the feature subset and tests the accuracy of classification algorithm built on the tentative feature subset. The feature that results in the lowest classification error is definitely added to the feature subset. After each addition step the algorithm removes any previously added feature if its removal decreases error. The process stops after a certain number of iterations provided by the user. Then the user determines the optimum features subset by examining the improvement in classification error with respect to features added at each iteration. Please note that, once the optimum feature subset is determined, there is no need to repeat this feature selection step any more, unless a new training database is available or new features will be explored. Hence, this step does not limit automatization of the method proposed in this paper.

2.2.5. Classification

Classification stage is applied to discriminate true segmentations (true SC's) from false ones found by object segmentation step, hence it is a binary decision. The task of a classifier is to assign the object to a category using features. An ideal (omnipotent) classifier should not be in need of a sophisticated feature extractor (Duda et al., 2001). However, it is not possible to find an omnipotent classifier that can solve all real-world problems in reality. Therefore, researchers concentrate on the goal of finding task-specific classification algorithms that should be as robust and accurate as possible. This goal can only be accomplished by comparing performances of several classification algorithms for the same task, which is lacking in the area of SC recognition. The following supervised classification algorithms are tested in this work.

Linear discriminant classifier (LDC), for two-category case, tries to find a linear decision boundary that separates the feature space into two half-spaces by minimizing the criterion function

$$g(x) = w^t x + w_0 \tag{2}$$

k-Nearest neighbor (k-NN) is one of the most popular classifiers used for classification. It assigns an object to the category, which is most represented among the k nearest samples of that object. Nearest samples are found by Euclidean distance measure. Performance of k-NN highly depends on the k parameter and size of training set, both of which should be large enough for optimal performance and small enough to avoid computational overload (van der Heiden, Duin, de Ridder, & Tax, 2004).

While assigning a new sample, k-NN classifier gives equal importance to the k nearest samples by assuming that they are equidistant from the new sample. However, using the distance information of the nearest samples by the following formula can improve the performance of conventional k-NN:

$$u_i(x) = \frac{\sum_{j=1}^{k} u_{ij}(\|x - x_j\|)^{-2/(m-1)}}{\sum_{j=1}^{k}(\|x - x_j\|)^{-2/(m-1)}} \tag{3}$$

where $u_i(x)$ is the predicted membership value of test sample x for class i, u_{ij} is the membership (either 0 or 1) of jth neighbor to the ith class and m is the *fuzzifier* parameter (set to 2) that determines how heavily the distance is weighted. This *fuzzified* classifier is known to be fuzzy k-NN.

The aim of *boosting* is to improve the accuracy of any given learning algorithm. *AdaBoost* (Rätsch, Onoda, & Müller, 2000) is the most popular one among the boosting methods. It is a meta learning algorithm for constructing a 'strong' learner (g) from linear combination of 'weak' learners (h_t). Thus, the classification decision for a test sample x is taken by

$$g(x) = \mathrm{sgn}\left(\sum_{t=1}^{t_{\max}} \beta_t h_t(x)\right) \tag{4}$$

where sgn(·) refers to signum function ($\mathrm{sgn}(v) = +1$ if $v > 0$, 0 if $v = 0$ and -1 if $v < 0$), β_t are the coefficients found by boosting process and $t_{\max}$ is the number of weak learners.

Support vector machines (SVM) is a statistical learning method based on structural risk minimization procedure (Burges, 1998). In the binary case, SVM classifier maps the input space into a new space through kernels and then tries to find the hyperplane that separates the classes with maximum margin in this new space. For a test sample x, its output class y is found by SVM as

$$y = \mathrm{sgn}\left(\sum_{i=1}^{N} \alpha_i y_j \mathbb{K}(x_i, x)\right) \tag{5}$$

where sgn(·) refers to signum function, N is the number of training samples, x_i is the ith training sample with $y_i \in -1, +1$ being its corresponding class label and $\mathbb{K}(x_i, x)$ is the kernel function. $\alpha = \alpha_1, \alpha_2, \ldots, \alpha_i, \ldots, \alpha_n$ are the Lagrangian multipliers found by maximizing the quadratic equation

$$L_{\mathrm{D}} = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j \mathbb{K}(x_i, x_j) \tag{6}$$

with $0 \leqslant \alpha_i \leqslant C\ \forall i$. C is the trade-off parameter between large margin and low margin classification error. x_i's for which $\alpha_i > 0$ are the support vectors. Three different kernel functions are tested with SVM: polynomial (Eq. 7), homogeneous polynomial (Eq. 8), and Gaussian radial basis function (Eq. 9). d refers to the degree of polynomial kernels, '·' denotes scalar product of its arguments and γ is the width of Gaussian kernel:

$$\mathbb{K}(x_i, x_j) = (x_i \cdot x_j + 1)^d \tag{7}$$

$$\mathbb{K}(x_i, x_j) = (x_i \cdot x_j)^d \tag{8}$$

$$\mathbb{K}(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{2\gamma^2}} \tag{9}$$

Performance estimation of the classification process is measured by K-fold cross-validation method, which works as follows: First dataset is partitioned into K non-overlapping subsets. Then each subset is used for testing, while remaining $K-1$ ones are used for training. classification error is the average error rate of K tests. Advantage of K-fold cross-validation is that all samples of dataset are eventually used for both training and testing. A high K value will result in a very accurate estimator, but computational expense will also be high. So, $K = 5$ is chosen for the tests done in this work. Furthermore, samples of the dataset are randomly ordered before being introduced to the classification algorithm, to prevent biased recognition with respect to sample order.

In this research, libraries of Canu, Grandvalet, and Rakotomamonjy (2003) and Rätsch et al. (2000) are used for SVM and AdaBoost classifications, respectively. The proposed system is implemented under Matlab 6 (R12.1) environment (MathWorks, Inc., 1984) and tested on a Intel Pentium IV machine with 1.5 GHz CPU and 256 MB memory.

3. Results and discussion

Background removal and object segmentation processes are applied on each four filter images individually and visual observations showed that results of 750 nm filter image were visually equal or superior than those of others. Therefore, results of 750 nm filter image are manually classified as true (segmented object is from a SC) or false SC's (segmented object is not from any SC's). Table 2 shows result of this manual classification. A promising observation is that, none of the SC's in the images of stems or calyxes are missed by object segmentation step. However, a significant amount of false SC's from healthy and defected images are also segmented. Examples of segmentations can be seen in Fig. 6, where contours of segmented objects are displayed over original fruit images. As seen, segmentation results are encouraging. However, objects are sometimes partially segmented either due to high color variation (top-right image) or as a result of binary erosion (bottom-center image). There are also some false SC's like in the bottom-left and bottom-right examples, where the former is a healthy skin and the latter is a defected one. Therefore, results of object segmentation should further be refined by classification step to remove false SC's.

Table 2
Manual classification of segmented objects

	No. of segmented objects	
	True SC	False SC
Healthy view only	39	105
Stem in view	148	N/A
Calyx in view	145	N/A
Defect in view	95	150
Whole database	427	255

In order to find optimum parameters for classification algorithms, as an initial test, feature selection step is bypassed and all features are introduced to the classification algorithms. Several values for parameters are tested; such as $1 \leqslant k \leqslant 15$ for k-NN and fuzzy k-NN; $2 \leqslant t_{\max} \leqslant 50$ (number of weak learners), $e^{-1} \leqslant \lambda \leqslant e^{-12}$ (regularization parameter) and various number of iterations for AdaBoost; and three kernel functions with several parameters, $1 \leqslant C \leqslant \infty$ (upper-bound for Lagrangian multipliers) and $e^{-1} \leqslant \lambda \leqslant e^{-14}$ (conditioning parameter of quadratic programming method) for SVM.

Fig. 7 displays the results of these several tests as a function of true positive (tpr) and false positive rates (fpr). Diagonal between perfect recognition and perfect rejection is also displayed to ease comparison. Classification algorithm laying in the left-top most part of such a graph is said to be the best, because it presents the highest tpr and the lowest fpr. As observed, LDC exhibits low fpr together with relatively low tpr. Results of k-NN classifier are generally higher in tpr value than those of fuzzy k-NN. SVM and AdaBoost perform better than all others, where the former slightly outperforms the latter.

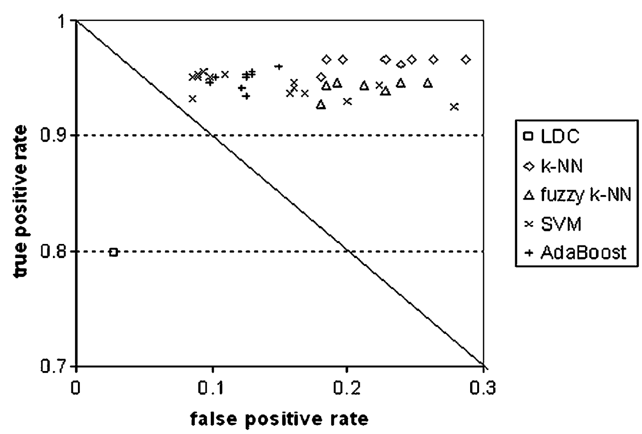

Fig. 7. Recognition performances of classification algorithms with various parameters.

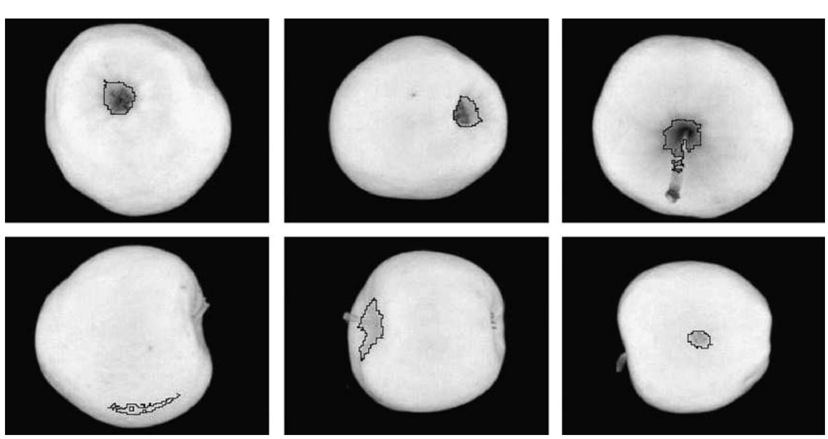

Fig. 6. Examples of object segmentation. Contours of segmented objects displayed over original images.

From this test optimum parameters for each classification algorithm are observed as; $k = 5$ for k-NN and fuzzy k-NN, $t_{\max} = 4$, $\lambda = 1e^{-6}$ and 10 iterations for AdaBoost and Gaussian RBF kernel with $\gamma = 5$, $C = \infty$ and $\lambda = 1e^{-2}$ for SVM. Next, classification algorithms with these best performing parameters are used together with SFFS method to examine if a small subset of features is advantageous in terms of classification accuracy. Our small database did not allow us to create an independent test set, hence testing is again performed by K-fold cross-validation method that permits separate training and testing subsets for each fold. Note that features of Fig. 5, extracted from each segmented object, are used once more in this test. Evaluation of overall recognition rate for each classification algorithm with number of features added is displayed in Fig. 8, where an exponentially decreasing increase is observed. After about eight features, they all reach to their plateau of highest recognition rate. Consistent with the previous results, algorithms can be sorted from worst to best as LDC, fuzzy k-NN, k-NN, AdaBoost and SVM, respectively. Highest recognition rate is observed by SVM with 9 features selected out of 35 and the corresponding detailed recognition results are presented in Table 3. Average, standard deviation, maximum, invariant moment and area features are found to be discriminative enough among the features set. In addition, absence of features measured from 500 nm filter image in the selected subset, makes us believe that 500 nm filter image is less important than others in SC recognition. 98% of tpr and 9% of fpr rates for whole database indicate that *type-I* (true SC's missed) and *type-II* (false SC's classified as SC) errors are 2% and 9%, respectively. This difference is even more significant as the number of false SC's are lower than that of true ones (see Table 2). This is probably because false SC's highly vary within each other. SC regions in the images of healthy database are correctly recognized with a rate of 85%. 99% and 100% of SC's in stem and calyx databases are correctly recognized, respectively, and only 13% of the defects were misclassified as SC. These rates show significant improvement with respect to the inspiring results of Kleynen et al. (2005), who stated 91%, 92%, and 17% rates for the same task and database, respectively. Despite these compelling improvements, pattern recognition methods are known to be training sample dependent. Hence, the proposed method should be further tested by a new set of test samples to acknowledge its performance.

Table 3
Recognition result of best feature subset of SFFS method for SC recognition by SVM

	tpr (%)	fpr (%)
Fruits with		
healthy view only	85	3
stem in view	99	N/A
calyx in view	100	N/A
defect in view	97	13
Whole database	98	9

Features: μ, 450 nm; σ, 450 nm; μ, 750 nm; max, 750 nm; σ, 750 nm; ϕ_1, 750 nm; max, 800 nm; σ, 800 nm; S

Features are from those displayed in Fig. 5. 'tpr' and 'fpr' refer to percentages of correctly found SC's and falsely classified objects that are not SC's, respectively.

Fig. 9 shows some of the objects misclassified by the proposed system. Our thorough observation of the missed objects revealed that

(1) if a true SC is located far from the center of fruit (close to the edges), then it has high probability to be misclassified (two top-left images);
(2) if two or more objects are touching, then they are likely to be misclassified (top-right image);
(3) if an object is partially segmented due to erosion, then it will most likely be incorrectly recognized (top-left image);
(4) low-contrast defects are most likely to be recognized as SC's (images below);
(5) *hail damage with perforation* and *frost damage* types of defects are among the most puzzling ones to discriminate from true SC's (two bottom-left images);
(6) *recent bruises* (1–2 h *old*), *russets* and *hail damages without perforation* are not confused with SC's at all.

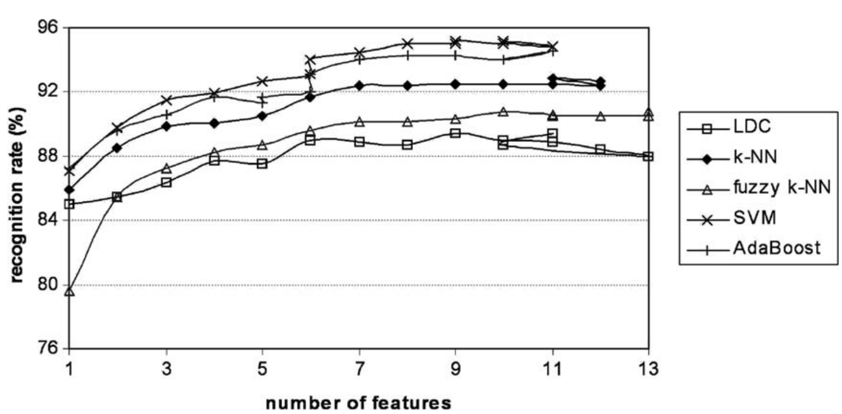

Fig. 8. Effect of feature selection on recognition rates of classification algorithms.

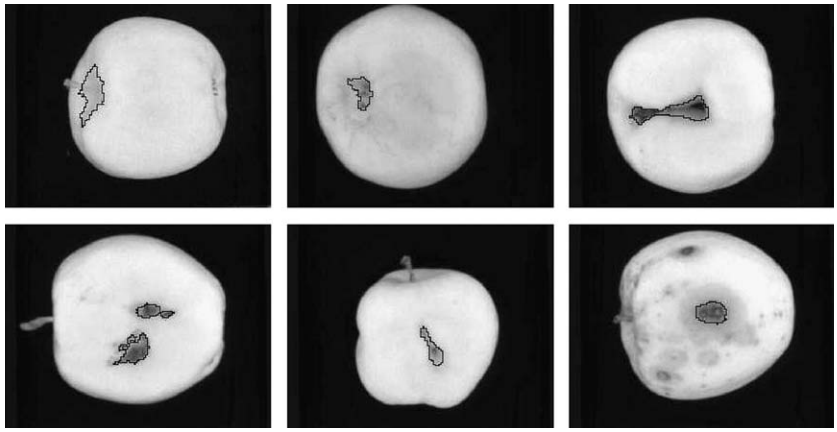

Fig. 9. Some misclassifications observed by the proposed approach.

4. Conclusion

External quality grading of apple fruits by machine vision is still an open, tedious and challenging problem. Accuracy of this task depends on several subtasks, one of which is precise recognition of stem, calyx areas.

In this paper we introduced a novel method to recognize stem, calyx regions in 'Jonagold' variety apples. The method consists of threshold-based segmentation, features extraction and classification algorithm-based recognition steps. 750 nm filter image is found to be the best for segmentation stage, which calculates statistical measures from the fruit area and then segments objects by adaptive thresholding. Statistical, textural and shape features are extracted from each segmented object and these features are fed to the classification algorithms after normalization step. Five supervised classification algorithms are tested in this work: LDC, *k*-NN, fuzzy *k*-NN, AdaBoost and SVM. SVM classifier gave the best performance in terms of true positive and false positive rates.

For the purpose of removing irrelevant or redundant features sequential floating forward selection method is tried with each classification algorithm to select the best-discriminating ones. Results showed that SVM is again the best choice among the algorithms tested. Furthermore, feature selection method removed 26 irrelevant/redundant features out of 35, resulting in a slightly higher recognition rate. Selected features reveal that *minimum*, *gradient*, *skewness*, *kurtosis*, *perimeter* and *circularity* features do not bring relevant information for classification. Moreover, features calculated from 500 nm filter image are not valuable enough to be selected. By the selected feature subset 99% of stems and 100% of calyxes are accurately recognized, and only 13% of defects are misclassified as SC.

Even though these encouraging results are obtained by separate training and test sets thanks to the *K*-fold cross-validation approach, performance of the SVM-based method has to be further tested by a new set of samples to verify that it is sample independent. Although the database is composed of images of stems and calyxes with various orientations relative to camera view, the tilting is not performed in a controlled manner. Thus, another future work is to prepare a new database and test the maximum orientation angle of stems and calyxes relative to camera axis before it will degrade the performance of SVM method.

Acknowledgements

This research is funded by the General Directorate of Technology, Research and Energy of the Walloon Region of Belgium with Convention No. 9813783. We would also like to thank Prof. M.-F. Destain, O. Kleynen and V. Leemans from Gembloux Agricultural University of Belgium for the image database. Finally, authors are especially grateful to the anonymous reviewers of this paper for their invaluable comments.

References

Bennedsen, B., & Peterson, D. (2004). Identification of apple stem and calyx using unsupervised feature extraction. *Transactions of ASAE, 47*(3), 889–894.

Burges, C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery, 2*, 121–127.

Canu, S., Grandvalet, Y., & Rakotomamonjy, A. (2003). Svm and kernel methods matlab toolbox. Perception Systémes et Information, INSA de Rouen, Rouen, France.

Cheng, X., Tao, Y., Chen, Y., & Luo, Y. (2003). NIR/MIR dual-sensor machine vision system for online apple stem-end/calyx recognition. *Transactions of ASAE, 46*(2), 551–558.

Crowe, T., & Delwiche, M. (1996a). Real-time defect detection in fruit—part I: Design concepts and development of prototype hardware. *Transactions of ASAE, 39*(6), 2299–2308.

Crowe, T., & Delwiche, M. (1996b). Real-time defect detection in fruit—part II: An algorithm and performance of a prototype system. *Transactions of ASAE, 39*(6), 2309–2317.

Duda, R., Hart, P., & Stork, D. (2001). *Pattern classification* (2nd ed.). New York: Wiley.

Hu, M. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory, IT-8*(February), 179–187.

Jain, A., & Zongker, D. (1997). Feature selection: Evaluation, application and small sample performance. *IEEE Transactions Pattern Analysis and Machine Intelligence, 19*(2), 153–158.

Kleynen, O., Leemans, V., & Destain, M.-F. (2003). Selection of the most efficient wavelength bands for 'jonagold' apple sorting. *Postharvest Biology and Technology, 30*(3), 221–232.

Kleynen, O., Leemans, V., & Destain, M.-F. (2005). Development of a multi-spectral vision system for the detection of defects on apples. *Journal of Food Engineering, 69*(1), 41–49.

Li, Q., Wang, M., & Gu, W. (2002). Computer vision based system for apple surface defect detection. *Computers and Electronics in Agriculture, 36*(2–3), 215–223.

MathWorks, Inc. (1984). Matlab: The language of technical computing. Available from http://www.mathworks.com.

Penman, D. (2001). Determination of stem and calyx location on apples using automatic visual inspection. *Computers and Electronics in Agriculture, 33*(1), 7–18.

Pudil, P., Ferri, F., Novovicova, J., & Kittler, J. (1994). Floating search methods for feature selection with nonmonotonic criterion functions. In *Proceedings of the 12th IEEE Conference on Pattern Recognition*, October 1994, Jerusalem, Israel (pp. 279–283).

Rätsch, G., Onoda, T., & Müller, K.-R. (2000). Soft margins for AdaBoost. *Machine Learning, 42*(3), 287–320.

Unay, D., & Gosselin, B. (2004). Stem-end/calyx detection in apple fruits: Comparison of feature selection methods and classifiers. In *Proceedings of the International Conference Computer Vision and Graphics (ICCVG)*, September 2004, Warsaw, Poland.

van der Heiden, F., Duin, R., de Ridder, D., & Tax, D. (2004). *Classification, parameter estimation and state estimation: An engineering approach using MATLAB*. New York: Wiley.

Wen, Z., & Tao, Y. (1999). Building a rule-based machine-vision system for defect inspection on apple sorting and packing lines. *Expert Systems with Applications, 16*(3), 307–313.

Wen, Z., & Tao, Y. (2000). Dual-camera NIR/MIR imaging for stem-end/calyx identification in apple defect sorting. *Transactions of ASAE, 43*(2), 449–452.

Yang, Q. (1996). Apple stem and calyx identification with machine vision. *Journal of Agricultural Engineering Research, 63*(3), 229–236.