

# Sensors for Fruit Firmness Assessment: Comparison and Fusion

V. Steinmetz; M. Crochon; V. Bellon Maurel; J. L. Garcia Fernandez; P. Barreiro Elorza;  
L. Verstreken

CEMAGREF, Agricultural Equipment and Food Process Engineering Division, BP 5095, 34033 Montpellier Cédex 1, France

Departamento de Ingeniería Rural, Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros Agrónomos, 28040 Madrid, Spain

Katholieke Universiteit Leuven, Department of Agricultural Engineering, Faculty of Agriculture, Kardinaal Mercierlaan 92, B-30001 Heverlee, Belgium

Non-destructive measurement of fruit firmness is a difficult problem and many different sensors have been developed in order to achieve this task. Three different European laboratories were associated in collaborative experiments on peaches, to compare three different sensing techniques, namely, sound, impact and micro-deformation. A Bayesian classifier is associated with each individual sensor and provides a classification into three categories, namely "soft", "half firm" and "firm". The fusion of the different sensors is performed by using Bayesian classifiers associated with heuristic methods for identity fusion. The result of the identity fusion is compared with the classification provided by an unsupervised algorithm based on destructive measurements. The fusion process provides some improvement in the classification results. For the individual sensors, the error rate of the classification varied from 19 to 28%, but the fusion process reduced this to 14%. Moreover, all measures of agreement between sensors lead to the conclusion that fusing sensors is better than using individual sensors.

## 1. Introduction

Peach quality is dependant on a combination of several factors such as firmness, sugar content, visual appearance and aroma. However, peach quality is highly related to firmness, as indicated by a Magness-Taylor firmness test (Rood<sup>1</sup>). Firmness is important because the peach has to be firm enough to minimize damage during packing and shipping. Moreover the stage of development of the peach has to ensure acceptable ripening at the terminal market. Colour is often accepted as the second index of maturity (Delwiche<sup>2</sup>).

Maturity sorting is a labour-intensive component of fresh packing lines, and non-destructive devices for on-line maturity assessment have been developed. Different techniques have been evaluated for non-destructive sensing of fruit firmness (sound, impact or microdeformation), but it still remains difficult to compare one sensor with another, or to determine the specific advantages of each sensor.

Sensor fusion is analogous to the cognitive process used by humans to integrate data continually from their senses to make inferences about the external world. So far it has been widely applied to military situations (battlefield surveillance, tactical situation assessment), or non-military purposes (robotics, automated manufacturing, remote sensing). Theoretical developments in sensor fusion have influenced our study on non-destructive firmness sensing, based on the fact that combinations of sensors should give a better result than each individual sensor alone.

Experimental sensors are supposed to measure the same property of the fruit, i.e. firmness. In order to compare them and to combine them, they have to be used on the same samples and in the same experimental conditions. A collaborative experiment was set up which enabled the comparison and the fusion of the firmness sensors. The object of the present work was (1) to compare the performance of each individual non-destructive sensor to the destructive sensors; (2) to apply sensor fusion techniques in order to determine peach firmness; (3) to find statistical tools in order to compare the fusion process with the individual sensors; and (4) to compare the performance of the fusion process with each individual sensor. The ultimate objective was to show whether a combination of sensors would give a better result than a single sensor.

## 2. Equipment and instrumentation

### 2.1. Sound based sensor

During the past two decades, frequency based measurement techniques have been studied for non-destructive evaluation of fruit firmness (Armstrong *et al.*<sup>3</sup>). The excitation is applied by striking the fruit at the top with a small hammer. A microphone, installed inside the support medium, but not in contact with the fruit, measures the vibration response signal. The measured signal is amplified in a signal conditioner, passes through a low-pass filter, and is finally sent to a computer, which extracts the frequency measurement. The raw output and the data processing for this sensor are presented in Fig. 1. The second resonant frequency  $f$  was shown to be related to apple firmness via the stiffness factor  $f^2m^{2/3}$ , where  $m$  is the apple mass. Features extracted from this curve by using the Hankel matrix-based total least square method (HTLS), lead to the resonant frequency  $f$ , and the stiffness factor  $f^2m^{2/3}$  (Chen and De Baerdemaeker<sup>4</sup>).

### 2.2. Micro-deformation based sensor

A non-destructive penetrometer has been developed by Cemagref in order to measure micro-deformation (Bellon *et al.*<sup>5</sup>). The principle of the measurement consists in applying a 4 mm diameter micro-sphere, which emerges from a reference surface, to the surface of the fruit with a constant force. The penetration of the sphere into the skin of the fruit is measured with high precision (300  $\mu\text{m}$  at maximum), and determines the fruit firmness. This penetration does not leave any permanent deformation on the surface of the fruit because of the small size of the microsphere. The apparatus is shown in Fig. 2. The result of the measurement is an index ranging from 0

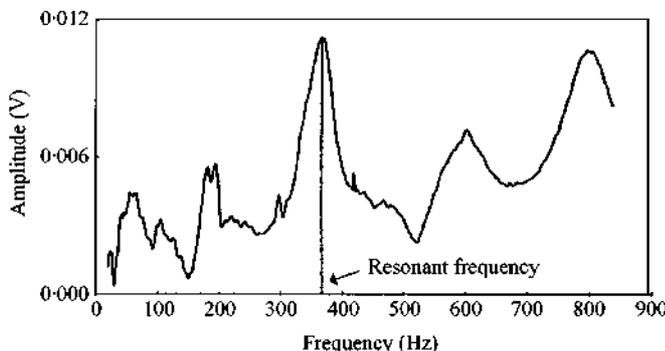


Fig. 1. Vibration response spectrum of a peach measured using a microphone

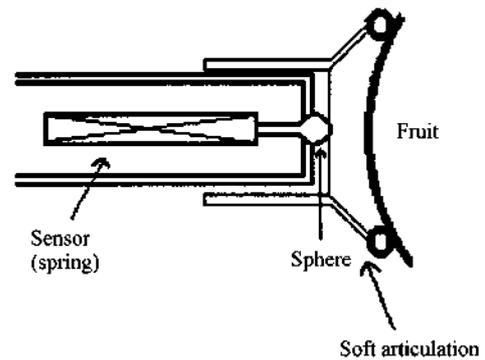


Fig. 2. Microdeformation based firmness sensor

to 100 (the apparatus is calibrated with a maximum deformation related to an index value of 100 representing a firm surface).

### 2.3. Impact response based sensor

Determination of peach firmness has been widely applied based on impact response (Mohsenin,<sup>6</sup> Delwiche,<sup>7</sup> Brusewitz<sup>8</sup>). The basic principles of this sensor are presented in Fig. 3. A non-destructive penetrometer using these principles has been developed (Ruiz<sup>9</sup>) to predict firmness. A 50 g steel sphere, instrumented with a miniature accelerometer, with a radius of curvature of 98 mm is dropped from a predetermined height depending on the fruit tested (4 cm, for peaches) onto the cheek of the fruit. The deceleration of the impactor during contact is measured versus time. The maximum deceleration corresponds to the point of maximum deformation. From the original response curve of deceleration versus time, four other parameters can also be extracted, namely, maximum deformation during impact, duration of impact, absorbed energy during impact and elastic modulus (maximum force/maximum deformation). The user of the sensor can select and extract the needed parameters. For peaches, the maximum force is the primary impact response parameter of interest.

### 2.4. Destructive sensors

Two different destructive sensors are used: a Magness-Taylor, which measures the maximum force, i.e. firmness, and an Instron-type machine, which measures force versus deformation, i.e. stiffness. The Magness-Taylor sensor consists of a plunger (23 mm length) attached to a calibrated spring scale which is graduated in kg. The 8 mm diameter round tip of the

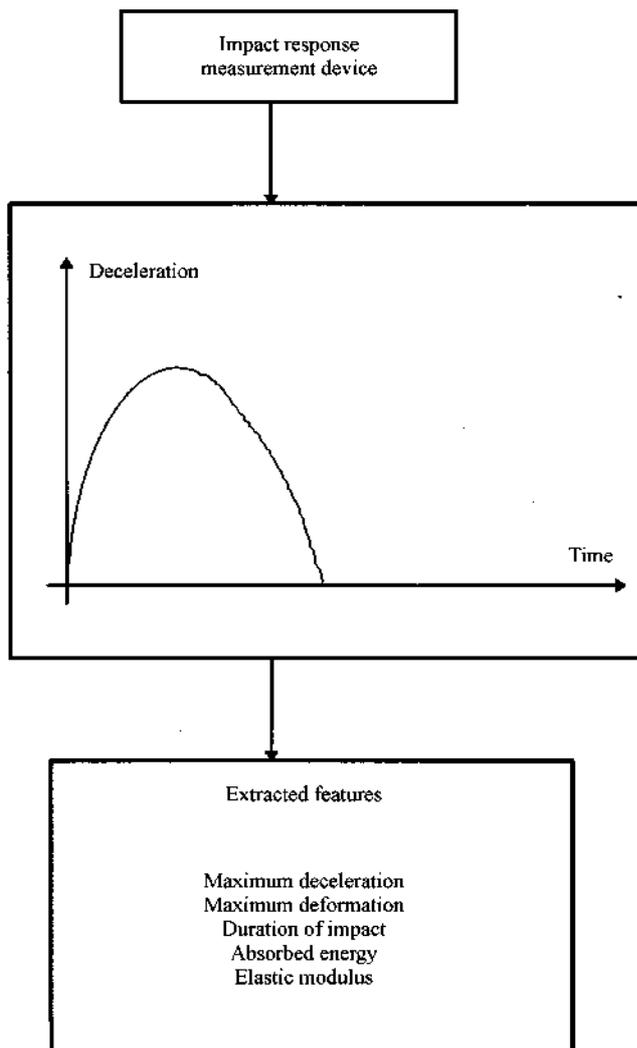


Fig. 3. Impact based firmness sensor

plunger is pressed with increasing force into the fruit to a depth of 8 mm, marked on the plunger. The penetrating force is then read on the scale. The Instron-type machine also uses a 8 mm steel plunger which is pressed into the fruit at a rate of 0.3 mm/s and the resulting force–deformation curve is recorded. The resulting curve displays a straight portion, for which force increases with deformation until the yield point is reached at which point the plunger breaks the skin and a sudden decrease in force is observed. The slope of the straight portion is recorded as a measure of stiffness. These two destructive methods are used in order to provide reference measurements.

These sensors provide information on the firmness of a small part of the fruit. However, the sound-based sensor provides a global assessment concerning firmness, while the impact and micro-deformation based

sensors provide a local assessment. One of the aims of this paper is to take advantage of this complementarity, and to combine these two types of measurements (local and global) to improve the firmness assessment.

### 3. Sensor comparison techniques and sensor fusion

#### 3.1. Measures of agreement between sensors

Before performing sensor fusion, it is important to check whether or not the sensors provide the same information. When the sensors are supposed to measure different properties from a sample, the sensors are complementary. In the opposite case, some of the sensors may be redundant, and the failure of one can be tolerated. Since three different sensors are used, it is interesting to compare the measurement of each individual sensor (variable  $x_i$ ) with the measurements given by the destructive sensors (variable  $y$ ). The literature provides some measures of association among variables in order to characterize the relationship with some numerical similarity measurements.

##### 3.1.1. Box-plot

The box-plot (Tukey<sup>10</sup>) is a graphical display that allows visual presentation of information extracted about properties (mean, variability and symmetry) of a set of data. This plot cannot be used for direct visual comparison of the means when the variables do not have the same unit, but it is still helpful for comparing variability and symmetry using reduced variables such

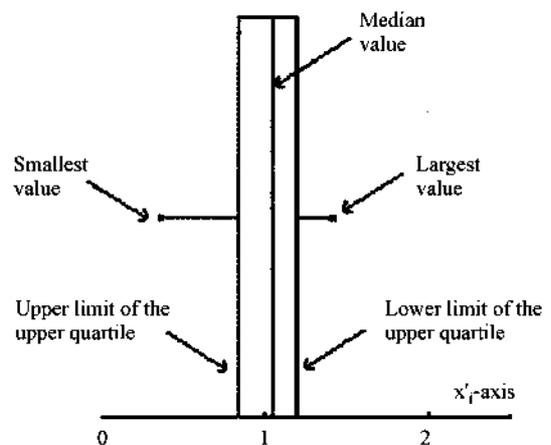


Fig. 4. Box and whisker plot

as  $x'_i = x_i/m$ , where  $m = \frac{1}{n} \sum_{i=1}^n x_i$  and  $n$  is the sample number. Fig. 4 shows an example of a box-plot diagram. This plot encloses the interquartile range of the reduced data in a box that has the median display within. The interquartile range has as its extremes the 75th percentile (upper quartile) and the 25th percentile (lower quartile). In addition, "whiskers" extend, showing extreme observations in the sample. The y axis does not provide any information, it is just the way the box-plot is usually sketched.

### 3.1.2. Correlation coefficient

The correlation coefficient measures the strength of a linear relationship between two variables. However the nature of the relationship is not known from the correlation coefficient itself, but some tests can be made to find this relationship. As a first step, the test for linearity of regression between  $x_i$  and  $y$  can be made, and a confidence interval (90%) can be computed for the correlation coefficient. The correlation coefficient provides a means of checking whether there is redundancy or complementarity between two sensors. Let  $\rho_{AB}$  be the correlation coefficient computed between the data measured with sensor A and the data measured with sensor B. Similarly, let  $\rho_{AC}$  be the correlation coefficient computed between the data measured with sensor A and the data measured with sensor C. If the confidence interval of the correlation coefficient  $\rho_{AB}$  overlaps the confidence interval of the correlation coefficient  $\rho_{AC}$ , then sensors B and C are said to be redundant with respect to sensor A, otherwise they are complementary.

The two previous indices are calculated based on the raw data provided by each sensor. However, when the sensors provide a classification, it is necessary to use other indices to measure the agreement between sensors. In this case, the results for these indices are dependent on the classification

process. This will be developed in the following sections.

### 3.1.3. $\chi^2$ values

Table 1 shows results of comparing the classification provided by a non-destructive sensor and the classification provided by a destructive sensor, where  $n$  is the number of samples used during the experiment,  $q$  is the number of classes,  $n_{i+}$  is the summation of the number of fruits on line  $i$ , and  $n_{+j}$  is the summation of the number of fruits on column  $j$ . Based on Table 1, a  $\chi^2$  value can be computed as defined in Eqn (1).

$$\chi^2 = \sum_{i=1}^q \sum_{j=1}^q \left[ \frac{\left( n_{ij} - \frac{n_{i+}n_{+j}}{n} \right)^2}{\frac{n_{i+}n_{+j}}{n}} \right] \quad (1)$$

When the classification of the two sensors totally agree, Table 1 results in a diagonal matrix, and  $n_{ij}$  equals zero when  $i$  is different from  $j$ . In that case, and by using Eqn (1), it can be shown that the value of  $\chi^2$  is equal to  $n(q-1)$ . For each non-destructive sensor, a  $\chi^2$  value can be computed and all the values are then sorted: the largest  $\chi^2$  value should indicate the non-destructive sensor that is most closely related to the destructive sensor.

### 3.1.4. Coefficient of contingency

The coefficient of contingency  $C$  (Maung<sup>11</sup>) can be computed as described in Eqn (2), where  $n$  is the sample number and  $q$  the class number.

$$C = \sqrt{\frac{\chi^2}{n(q-1)}} \quad (2)$$

It can be shown that the value of  $C$  belongs to the interval  $[0, 1]$ , and  $C = 1$  when the classification results in a diagonal matrix because  $\chi^2 = (q-1)n$  as derived in the previous section. The coefficient  $C$  should

**Table 1**  
**Two-sensor classification**

		Classification made by the destructive sensor				
		Class 1	Class 2	...	Class $q$	Summation
Classification made by a non destructive sensor	Class 1	$n_{11}$	$n_{12}$	...	$n_{1q}$	$n_{1+}$
	Class 2	$n_{21}$	$n_{22}$	...	$n_{2q}$	$n_{2+}$
	...	...	...	$p_{ij}$	...	$n_{i+}$
	Class $q$	$n_{q1}$	$n_{q2}$	...	$n_{qq}$	$n_{q+}$
	Summation	$n_{+1}$	$n_{+2}$	$n_{+i}$	$n_{+q}$	$n$

$n_{ij}$  = number of fruits classified into class  $i$  by the non-destructive sensor while they are classified into class  $j$  by the destructive sensors.

provide a measure of association between the classification made by the sensors. However, some authors (Goodman and Kruskal<sup>12</sup>) have pinpointed the major problem, that is to say, the difficulty in the interpretability of  $C$ . Alternative measures have been proposed based on the ranking process and on optimal class prediction.

### 3.1.5. Similarity index

Cohen<sup>13</sup> proposed a similarity index for two sensors A and B which independently classifies items into one of  $q$  mutually exclusive categories. This similarity measure is computed with the probabilities  $p_{ij}$  that sensor A classifies an object into category  $i$  while sensor B classifies the same object into category  $j$  (Table 2). The probability that sensor A will classify the observations into category  $i$  regardless of sensor's B classification of the observations into the categories, is then given by Eqn (3). Similarly, the probability that sensor B will classify the observations into category  $i$  regardless of sensor's A classification of the observations into the categories, is given by Eqn (4).

$$p_{i+} = \sum_{j=1}^q p_{ij} \quad (3)$$

$$p_{+i} = \sum_{j=1}^q p_{ji} \quad (4)$$

The proportion of cases where the sensors agree is defined as  $\theta_1$  in Eqn (5). These are two reasons for which the sensors can agree on the classification. The first one is chance, and the proportion of cases where the sensors agree purely by chance is defined as  $\theta_2$  in Eqn (6). The second one is the ability of the two sensors to provide the same classification because they measure the same property. This reason can be named "determinism" because it is the way the sensors were built or used for this purpose.  $\theta_1 - \theta_2$  is then the proportion of cases where the two sensors agree by determinism, and the similarity index is defined as  $K$

in Eqn (7). The similarity index is especially appropriated for independent sensors performing identity declarations. For complete agreement of the two sensors,  $\theta_1 = K = 1$ , and  $\theta_2 = 0$ .

$$\theta_1 = \sum_{i=1}^q p_{ii} \quad (5)$$

$$\theta_2 = \sum_{i=1}^q (p_{i+})(p_{+i}) \quad (6)$$

$$K = \frac{\theta_1 - \theta_2}{1 - \theta_2} \quad (7)$$

$\theta_1$  and  $\theta_2$  are calculated based on the probabilities  $p_{ij}$ . Probabilities  $p_{ij}$  are in general unknown, and are often replaced by estimated or observed values as in Eqn (8).

$$p_{ij} = \frac{n_{i+}n_{+j}}{n} \quad (8)$$

### 3.2. Review of sensor fusion techniques

Multisensor fusion is an evolving technology concerned with the problem of how to combine data from multiple sensors in order to make inferences about a physical event, activity, or situation (Hall and Nauda<sup>14</sup>). The ultimate objective of multisensor fusion is to find a combined declaration based on two (or more) sensors. The approach relies on the hypothesis that fusion of signals from several sensors (similar or disparate sensors) will give better results than the signal from one single sensor (Holmbom *et al.*<sup>15</sup>).

Data fusion is composed of three different categories, i.e. techniques for positional fusion, methods for identity fusion and ancillary support algorithms,

**Table 2**  
Two-sensor classification of probabilities

		Classification made by sensor B				
		Class 1	Class 2	...	Class q	Summation
Classification made by sensor A	Class 1	$p_{11}$	$p_{12}$	...	$p_{1q}$	$p_{1+}$
	Class 2	$p_{21}$	$p_{22}$	...	$p_{2q}$	$p_{2+}$
	...	...	...	$p_{ij}$	...	$p_{i+}$
	Class q	$p_{q1}$	$p_{q2}$	...	$p_{qq}$	$p_{q+}$
	Summation	$p_{+1}$	$p_{+2}$	$p_{+j}$	$p_{+q}$	$p_{++}$

$p_{ij}$  = probability that sensor A classifies an object into category  $i$  while sensor B classifies the same object into category  $j$ .

i.e. data pre-processing techniques. The first category has been mostly applied to military situations (target tracking), and the third category to data alignment and pre-processing techniques. Since pre-processing techniques have already been developed for each individual sensor, this paper will focus on the application of the identity fusion techniques. Fusion of identity declarations can be done at different levels: a raw data-level (i.e. frequency spectrum), a feature level (i.e. stiffness factor), a decision-level (i.e. the fruit firmness prediction) or a combination of the three previous methods.

### 3.3. Fusion identity declaration

Since feature-level extraction has already been developed for each individual sensor, the decision-level fusion was applied to the identity declaration level of each individual sensor (Fig. 5). Statistical classification includes the Bayes minimum risk classifier (Tou and Gonzalez<sup>16</sup>). This classifier was applied to each non-destructive sensor for the firmness evaluation, and for separating the peaches into the three categories commonly used by the growers and called firm, soft and half firm.

Decision level fusion seeks to process identity declarations from multiple sensors to achieve a joint declaration of identity. Techniques for decision-level fusion have been described (Hall<sup>17</sup>) and include (1) classical inference, (2) Bayesian inference, (3)

Dempster-Shafer (D-S) method, (4) generalized evidence processing (GEP) theory and (5) heuristic methods.

The performance of various identity fusion techniques has not been studied in a systematic way. Several authors have performed comparisons under limited circumstances (Abdulghafour and Abidi<sup>18</sup>). Unfortunately, numerical comparisons of D-S, GEP and Bayesian techniques are very limited and more research needs to be performed in this area to assess the comparative inference accuracy for these alternative techniques. Therefore, heuristic methods were used in order to treat the identity fusion problem as if a group of humans were faced with a decision problem.

Applicable heuristic methods of identity data fusion include voting methods, scoring models, ordinal ranking techniques, Q-sort methods and pairwise ranking. Voting methods address the identity fusion problem by a democratic process. The decisions from  $N$  sensors are simply counted as votes with a majority or plurality decision rule. The decision process can be refined by giving a weighting for the specific advantages of each individual sensor. A "confidence" factor  $c_{ij}$  can be defined for sensor  $i$  relative to the classification of the samples into class  $j$ . For example, if sensor  $i$  is known to be efficient for some classification aspects (no errors are made by this sensor when a sample is classified into class  $j$ ),  $c_{ij} = 1$ . These coefficients can be defined based on the *a priori* knowledge of the performance of each individual sensor. If  $c_{ij} = 1$ , the sample is classified into class  $j$

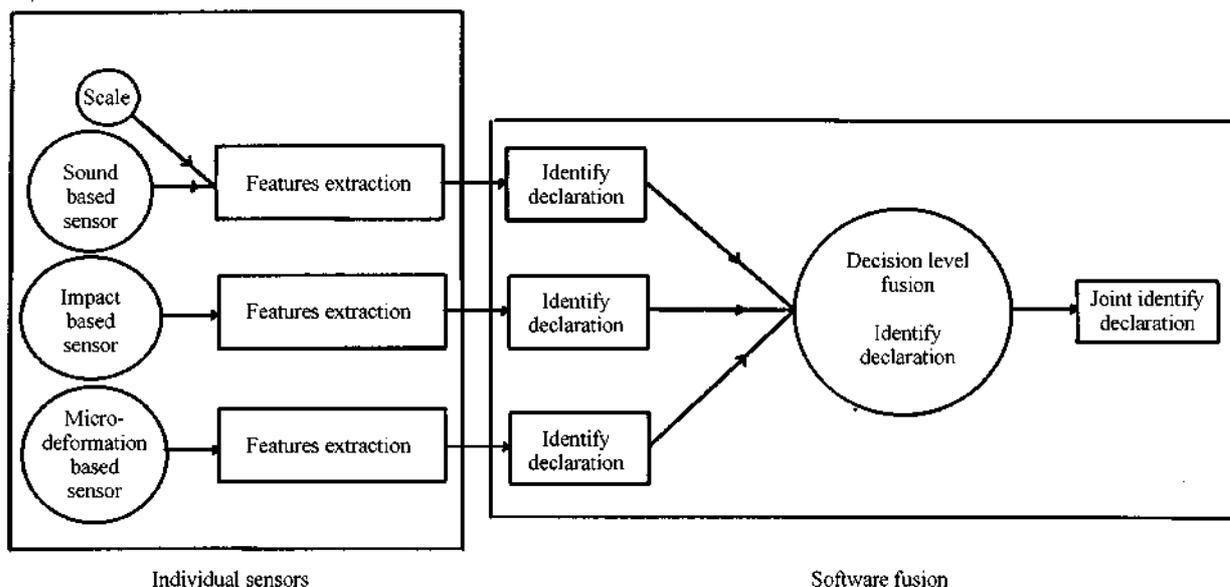


Fig. 5. Architecture for sensor fusion

and no voting method is applied. Otherwise, a majority voting method is applied.

#### 4. Experimental procedures and data acquisition

##### 4.1. Experimental procedures

Three different varieties of peach (Early Red Haven, 95 samples; Royal Glory, 69 samples; and Pavie, 39 samples) representing a total of 203 fruits were collected in early July 1993 from a grower located in the Rhône Valley (France). These different varieties present different levels of firmness: the variety "Pavie" is generally firmer than the variety "Royal Glory", which is firmer than the variety of "Early Red Haven".

This *a priori* knowledge was used in order to build the Bayesian classifier. *A priori* probabilities were set as indicated in Table 3. These values were chosen based on the knowledge of the properties of each variety, and the ratios were then arbitrarily chosen with the help of an expert in plant physiology in order to take into account this knowledge.

The samples were collected at different harvesting dates. Each sample was processed successively through the three individual sensors for non-destructive measurements in the following order: sound-based sensor, micro-deformation based sensor, and impact based sensor. Destructive measurements were then made with the destructive sensors, i.e. Magness-Taylor and an Instron-type machine.

Among the 203 samples, about one third were used to build the Bayes classifier, and the other two thirds were used to test the classifier. For the manual penetrometer and the impact sensor, to measurements were made on each peach: the first one on the "greenest" part of the peach, the second one on the opposite part. For the micro-deformation sensor, three measurements were made on each side of the fruit, and for the sound sensor, three measurements were made on each fruit. For the Instron-type machine, one measurement was made. Data were then averaged for each sensor, and the average of the

measurements was supposed to represent peach firmness as a whole.

##### 4.2. Reference measurements

The destructive measurements were used as references, and were classified using an iterative partitioning method (centre mobile algorithms, Anderberg<sup>19</sup>). Data were classified into three classes, corresponding to the three classes "soft", "firm" and "half firm". The algorithm was applied once to the data, and repeated 20 times (Roux<sup>20</sup>). Different classifications are obtained because the algorithm is initialized with a random process, i.e. the three initial cluster centres are randomly initialized. Samples belonging 20 times to the same class are allocated to this class. This leads to the three classes "firm", "soft", "half firm". Other samples that do not belong to these classes are classified into two fuzzy classes: "half firm or soft" for the samples that belong sometimes to the "soft" or sometimes to the "half-firm" class, and "half firm or firm" for the samples that belong sometimes to the "firm" or sometimes the "half-firm" class. This process resulted in five classes.

##### 4.3. Sensor data

Data were acquired from each individual sensor. The type of output values are presented in Table 4.

#### 5. Results and discussion

##### 5.1. Sensor similarities based on raw data

###### 5.1.1. Box-plot

Fig. 6 is a presentation of the various box-plot diagrams made with the 203 fruits. There is greater similarity between Fig. 6(d) and (c) than between Fig. 6(d) and (a) or (b). This suggests that the micro-deformation sensor shows the greatest similarity to the reference (Magness-Taylor sensor). However, the median values of these two diagrams are different. The width of Fig. 6(a) is small because the standard deviation of the measurements provided by the impact sensor is 0.24, which is relatively low compared with the standard deviation of the other sensors (0.43 for the sound sensor, 0.70 for the micro-deformation sensor and 0.69 for the Magness-Taylor sensor).

**Table 3**  
*A priori probabilities versus peach variety*

Quality/Variety	Soft	Half-firm	Firm
Early Red Haven	0.33	0.33	0.33
Royal Glory	0.3	0.3	0.4
Pavie	0.25	0.25	0.5

**Table 4**  
**Data acquisition structure**

<i>Sensor type</i>	<i>Sensors</i>	<i>Features</i>	<i>Units</i>
Non-destructive	Sound	Stiffness factor	$\text{g}^{2/3} \text{s}^{-2}$
	Impact	Maximum force	N
Destructive	Micro deformation	Firmness index	%
	Magness–Taylor	Maximum pressure	$\text{kg}/0.5 \text{ cm}^2$
	Instron	Slope	$\text{kg}/\text{mm}$

### 5.1.2. Correlation coefficients between the non-destructive and destructive sensors

Table 5 shows the correlation coefficient values and their confidence intervals. For a linear regression model between non-destructive and destructive measurements, the confidence intervals of the correlation coefficients from sensor “impact” and “microdeformation” do overlap. In other words, the percentage of variation in the firmness measurements do not change significantly (at a 90% level) from sensor “impact” to sensor “micro-deformation” for a linear relationship. This tends to prove that the sensors based on impact and microdeformation are redundant, since the sensor based on sound is complementary. The correlation coefficient found for the sensor “sound” is somewhat different from the other, which is an expected result. As stated in Section 2.4, the destructive sensors and the impact and microdeformation based sensors provide a local assessment of firmness, since the sound based sensor provides a global assessment of firmness.

Table 5 shows also that there is a better correlation between the different non-destructive sensors than between the destructive sensors as a reference.

## 5.2. Sensor fusion

### 5.2.1. Unsupervised classification based on destructive sensors

The result for unsupervised classification based on the Magness–Taylor and Instron-type sensors is presented in Table 6. The decision boundaries used by experts for classification of peach firmness are presented in Table 7. The boundaries listed in Table 7 are quite similar to the ones provided by the unsupervised classification, because the mean values of the two fuzzy classes “half-firm or soft” and “half-firm or firm” are more or less similar to the limits used by the expert. Unsupervised classification is not only a confirmation of the experts classification, but it also allows classes to be defined without firm limits. These fuzzy classes allow inclusion of samples

for which it is not easy to determine the correct classification between “half firm” and “firm” or “half firm” and “soft”.

Table 8 describes the number of fruits used to build the Bayes classifier for each class. The fruits allocated to the fuzzy classes were the results of the centre mobile algorithms described in Section 4.2. The other fruits were randomly put into the training and into the test set, so that in total, one third of the fruits are in the training set, and two thirds are in the test set. Samples from the fuzzy classes were not used in the training set because there is a high level of uncertainty concerning their belonging to one class. However, these samples were used in the test set, and were considered as samples belonging to the contiguous classes.

### 5.2.2. Supervised classification for the non-destructive sensors

Tables 9 to 11 present the number of fruits resulting from the classification of the different non-destructive sensors with a Bayes classifier. The performance of the classification is evaluated by computing the number of errors. As a general rule, an error occurs when the assigned class provided by the non-destructive sensor is different from the class assigned by the destructive sensors. However, this rule is not applied for the fuzzy classes (“soft or half-firm” and “half-firm or firm”), because a sample that belongs to a fuzzy class can be assigned to either one class or the other by the non-destructive sensor. For example, a sample assigned to the fuzzy class “soft or half-firm” by the destructive sensors can be assigned by the non-destructive sensor to the class “soft” or the class “half-firm” without considering this as an error. But, if the destructive sensor assigns the sample to the class “firm”, it is an error. The percentage error is then defined as the total number of errors divided by the total number of samples. It can be seen from these tables that the impact sensor does not provide any classification errors for classes “soft” and “firm”, and the micro-deformation sensor does not provide any classification errors for the class “firm”.

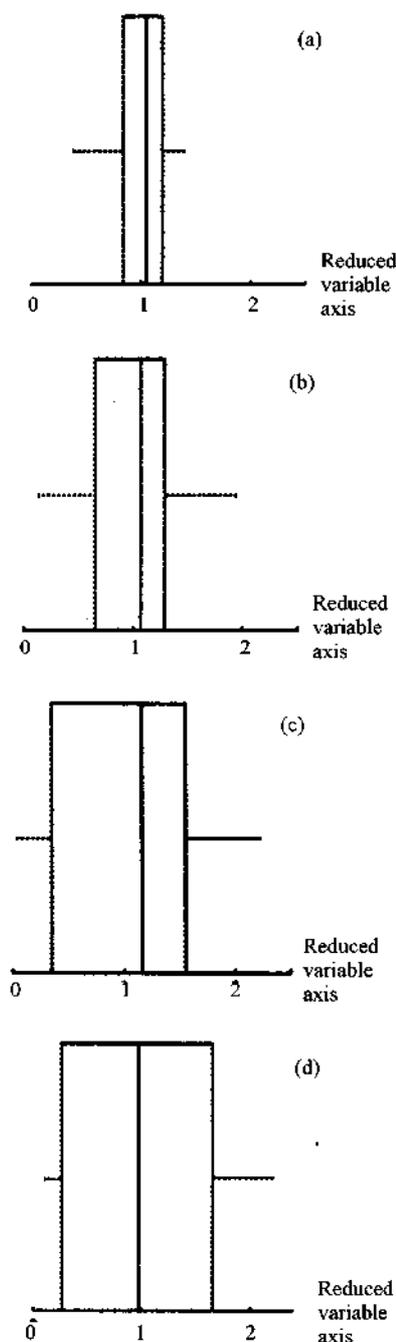


Fig. 6. Box-plots for the different sensors. (a) Impact; (b) sound; (c) microdeformation; (d) destructive sensor (Magness-Taylor)

### 5.2.3. Fusion classification

Based on the symbols defined in Section 3.3, the decision at the fusion level was made with the following rules.

(1) From Table 10, it can be stated that the impact based sensor is efficient for classifying soft and firm fruits, therefore the confidence factors for the impact

based sensor are chosen to be  $c_{\text{impact,soft}} = 1$  and  $c_{\text{impact,firm}} = 1$ . Similarly, from Table 11, it can be stated that the micro-deformation based sensor is efficient for classifying firm fruits, therefore the confidence factors for this sensor is chosen to be  $c_{\text{microdeformation,firm}} = 1$ . There is of course a possibility of conflict between sensor "impact" and sensor "microdeformation". This case of conflict is solved by rule number 2.

(2) When the confidence factor is not equal to 1, or when a conflict is detected, a majority voting method is applied.

Results of fusion classification are presented in Table 12. These results are obtained with a Bayesian classifier associated with each individual sensor, and the fusion classification method described in the previous paragraph. For example, a fruit can be classified as soft by the Bayesian classifier associated with the micro-deformation based sensor, as firm by the Bayesian classifier associated with the impact based sensor, and as soft by the Bayesian classifier associated with the sound based sensor. Since  $c_{\text{impact,soft}} = 1$  and  $c_{\text{microdeformation,firm}} = 1$ , there is a conflict between the two sensors "impact" and "microdeformation". Then the majority voting method is applied: two sensors classify the fruit as "soft", which is the decision of the fusion process. Results of Table 12 have to be compared with the classification given by each individual sensor. The fusion process slightly decreases the classification error rate (from 19% down to 14%). It is important to realise that the error rate could not be less than 8% with this fusion method because all the sensors disagree in the same way for 10 out of the 129 samples. In other words, the classification of the samples into class "half-firm", when they actually belong to class "soft" or "firm", could not be improved because the three sensors classified these samples in the class "half-firm".

One must try to explain the error rate of the fusion classification. First of all, only one feature (maximal force) was used from the impact sensor. The other features (maximum deformation, duration of impact, absorbed energy and elastic modulus (maximum force/maximum deformation) could have been used for this sensor. Secondly, it must be called that a few measurements were made for each fruit, and these data were then averaged. This can be a source of error for the classification, because averaging leads to a loss of information. Data could have been processed by using the two measurements made on each side of the peach, but this does not match the industrial objective of the non-destructive sensor, that is to say to provide a firmness assessment on the fruit as a whole, and not locally on a part of it.

**Table 5**  
Correlation coefficients and 90% confidence intervals between the sensors

	<i>Impact</i>	<i>Sound</i>	<i>Magness-Taylor</i>	<i>Instron</i>
Micro-deformation	0.93	0.82	0.88 [0.85–0.91]	0.80 [0.76–0.84]
Impact	1	0.86	0.84 [0.80–0.87]	0.79 [0.75–0.83]
Sound	—	0	0.67 [0.59–0.73]	0.64 [0.54–0.74]
Magness-Taylor	—	—	1	0.88 [0.85–0.91]

A source of error could also come from the destructive reference measurements. When all-non-destructive sensors agree, they may be right and the destructive measurement may be suspect, as has already been mentioned in Section 5.1.2.

Another source of error can be the small size of the training set (24 samples for the class “half firm” and 16 for the class “firm”). However, since unsupervised classification was used to build the classes based on destructive measurements, it is rather difficult to forecast in the experimental design how many samples will be used in the training set. Using supervised classification could be a way of avoiding this problem, but would remove all the advantages of unsupervised classification, i.e. the use of fuzzy classes.

It has been shown in Section 5.1.2 that the impact sensor is redundant with the microdeformation sensor when both are compared with the Magness-Taylor sensor. Table 13 describes the results when one of the sensors is removed from the fusion process. It should be pointed out that the voting method is not applicable when only two sensors are used. In that case, the decision is made by choosing the sensor for which the risk provided by its minimum risk classifier is smaller than the risk provided by the minimum risk classifier of the other sensor.

Table 13 shows that combining the sound sensor with the impact or the microdeformation sensor does

not change the classification error rate. However, combining the impact and the microdeformation sensors without the sound sensor provides a classification error rate which is slightly higher than the one provided by the fusion system with the three sensors. In order to obtain the best results, the three sensors should be included in the fusion process.

### 5.3. Sensor fusion evaluation

#### 5.3.1. $\chi^2$ test

Fig. 7 shows the different computed values of  $\chi^2$  for each classification sensor. These values were computed based on the results of the classification of each sensor (Tables 8 to 10), and based on the result of the fusion sensor classification (Table 11). Fig. 7 shows that the fusion process provides the highest score, i.e. the fusion process does improve the classification process.

#### 5.3.2. Coefficient of contingency

Fig. 8 shows the results of the computation of the  $C$  values for each sensor. These values are computed based on the  $\chi^2$  values from Fig. 8. First of all, the correlation coefficient for the fusion process is larger than any other coefficient. It tends to prove that the fusion process was efficient in improving the classification.

**Table 6**  
Results for unsupervised classification based on destructive measurements

	<i>Firmness classes</i>	<i>Soft</i>	<i>Half-firm or soft</i>	<i>Half-firm</i>	<i>Half-firm or firm</i>	<i>Firm</i>
	Number of fruits	90	8	46	22	37
Mean	Magness-Taylor	0.94	2.53	3.92	5.00	5.96
Standard deviation	kg/0.5 cm <sup>2</sup>	0.52	0.08	0.59	0.14	0.45
Maximum		2.25	2.65	4.75	5.20	6.80
Minimum		0.25	2.40	2.80	4.75	5.30
Mean	Instron	0.24	0.44	0.68	0.96	1.12
Standard deviation	kg/mm	0.11	0.11	0.28	0.27	0.23
Maximum		0.56	0.62	1.48	1.51	1.65
Minimum		0.06	0.03	0.05	0.09	0.43

**Table 7**  
Firmness classes used by experts

Quality	Soft	Half-firm	Firm
Magness–Taylor	<2	2–5	>5

Firmness measured with a manual penetrometer, peak pressure, kg/0.5 cm<sup>2</sup>.

**Table 8**  
Number of samples used for the classification

Quality	Soft	Half-firm	Firm	Total
Training set	34	24	16	74
Test set	56 + 8*	22 + 8 + 22*	21 + 22*	129
Total	90 + 8*	46 + 8* + 22*	37 + 22*	203

\* Samples belonging to the fuzzy classes

However, it must be emphasized that this correlation coefficient is dependant on the architecture of the fusion process. Choosing another architecture would have led to another value for the coefficient. From the C value, it is possible to say that there is a higher association with the fusion process than any other sensors, but the remark about difficulty of interpretation made at the end of Section 4.1 adds caution to this statement. However, it can be seen that the order of the C coefficients confirms the classification of sensors that was found with the correlation coefficient.

### 5.3.3. Similarity index between destructive and non-destructive sensors

The values of the similarity index are presented in Table 14. This table provides confirmation of what has already been found by using the  $\chi^2$  test and the coefficient of contingency. In other words, the order of the similarity index with respect to the different non-destructive sensor is the same as the order of the  $\chi^2$  values and the order of the coefficient of contingency. However, by using the coefficient of contingency, we have additional information concerning the proportion of cases in which the non-destructive sensors non-randomly agree with the destructive sensors. This proportion is around 36% for each sensor.

### 5.4. Discussion

The box-plot provides some graphical interpretation of the data, but this interpretation remains rather subjective. Correlation coefficients enable a check to

**Table 9**  
Classification for sound based sensor

Destructive sensor					
	Soft	Soft or half-firm	Half-firm	Half-firm or firm	Firm
<i>Non-destructive sensor</i>					
Soft	42	4	6	2	4
Half-firm	13	3	16	11	9
Firm	1	1	0	9	8

Total error: 28%.

**Table 10**  
Classification for impact based sensor (maximum force)

Destructive sensor					
	Soft	Soft or half-firm	Half-firm	Half-firm or firm	Firm
<i>Non-destructive sensor</i>					
Soft	41	1	0	0	0
Half-firm	15	7	22	16	16
Firm	0	0	0	6	5

Total error: 24%.

**Table 11**  
Classification for micro-deformation based sensor

Destructive sensor					
	Soft	Soft or half-firm	Half-firm	Half-firm or firm	Firm
<i>Non-destructive sensor</i>					
Soft	49	1	2	0	0
Half-firm	7	7	20	12	15
Firm	0	0	0	10	6

Total error: 19%.

**Table 12**  
Classification for the fusion process by using all non-destructive sensors (Sound-based sensor, micro-deformation based sensor and impact based sensor)

Destructive sensor					
	Soft	Soft or half-firm	Half-firm	Half-firm or firm	Firm
Soft	50	1	2	0	0
Half-firm	6	6	20	9	9
Firm	0	1	0	13	12

Total error: 14%.

**Table 13**  
**Classification results by removing one of the sensors**

Sensor combination	Error rate
Sound and impact	20%
Sound and microdeformation	19%
Impact and microdeformation	16%

be made on which sensors are redundant or complementary, and produce some questions about the validity of the destructive sensor measurements. The correlation coefficient is the most interesting parameter before performing the fusion process. Though the error rate was still 14%, the chosen architecture for sensor fusion was reasonably successful with regard to the industrial requirements. However, all the coefficients ( $\chi^2$  test, coefficient of contingency, similarity index) show that the fusion process was more efficient than any individual sensor. The coefficient of contingency is interesting because it provides the performance of the fusion on a normalised scale. The similarity index is also interesting, since it provides information on the agreement of sensors only by chance. Therefore, these two coefficients are the more attractive for evaluation of the fusion process.

The variety of the peaches appears to be an important factor in order to refine the classification during the classification at the decision level, as it influences the probability of occurrence of the firmness classes for each variety. *A priori* knowledge of the performance of each individual sensor turns out to be valuable information in the fusion process design.

## 6. Conclusion

In this paper a model was tested which offers the possibility of fusing different non-destructive firmness sensors. This model integrates the features extracted from each individual sensor, and associates them with a Bayesian classifier and a voting method that performs a joint identity declaration of peaches among

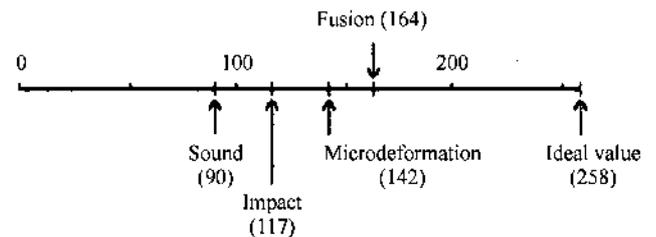


Fig. 7.  $\chi^2$  values for the different sensors

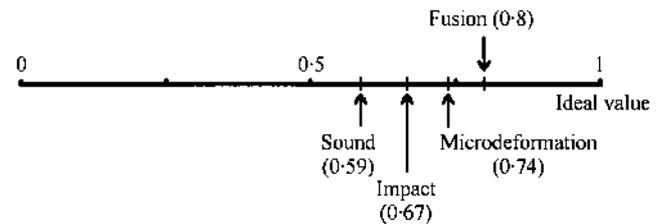


Fig. 8. Coefficient of contingency (*C*) values for the different sensors

three classes “soft”, “half firm” and “firm”. The three classes are provided by the mobile centre algorithms based on destructive sensors.

Different tools (box-plot) or indices (correlation coefficient,  $\chi^2$  test, coefficient of contingency, similarity index) were described in order to assess whether the fusion process does or does not improve the classification. It was shown that the coefficient of contingency and the similarity index are the more attractive for evaluation of the fusion process, but it is important to note that all indices led to the same conclusions, as there appears to be consistent results in the errors observed for firmness classification.

Only one fusion technique (applied at the decision level) was investigated in this study, and the fusion process provided some improvement in the classification results. The error rate of the classification decreased to 14% for the fusion process, while it varied from 19 to 28% for the individual sensors. Moreover, all measures of agreement between sensors lead to the conclusion that fusing of sensors is better than using individual sensors. More developments are on the way

**Table 14**  
**Similarity index between sensors**

	Similarity index	Sound	Impact	Microdeformation	Fusion
Magness-Taylor	$\theta_1$	0.72	0.76	0.81	0.86
	$\theta_2$	0.36	0.37	0.36	0.35
	<i>K</i>	0.56	0.62	0.71	0.79

for sensor fusion techniques applied to fruit quality assessment. Firmness prediction was the object of this paper, but fusion could also be performed on different kinds of sensors measuring, for example, sugar content, visual appearance, aromas and firmness. Also a wider number of parameters for each sensor could be used. This will be the object of future research.

### Acknowledgements

We express our appreciation to the EEC for their financial support of this research (Camar project, number 8001-CT91-0206) and to the Vivacoop growers in Aubenas, Ardèche, France for providing the experimental samples.

### References

- <sup>1</sup> **Rood P** Development and evaluation of objective maturity indices for California freestone peaches. Proceedings of the Journal of American Society of Horticultural Science 1957, **70**: 104–112
- <sup>2</sup> **Delwiche M J** Ground color as a peach maturity index. Journal of American Society of Horticultural Science 1981, **110**(1): 53–57
- <sup>3</sup> **Armstrong P; Zapp H R; Brown G K** Impulsive excitation of acoustic vibrations in apples for firmness determination. Transactions of the ASAE 1990, **33**(4): 1353–1359
- <sup>4</sup> **Chen H; De Baerdemaeker J** Total least square technique for estimating the vibration parameters of the apple from the time domain impulse response signal. Journal of Agricultural Engineering Research 1995, **61**(4): 283–290
- <sup>5</sup> **Bellon V; Vigneau J L; Crochon M** Non-destructive sensing of peach firmness. Proceedings of the fourth international symposium on fruit, nut and vegetable production engineering, Valencia, Spain, 1993, (2): 291–297
- <sup>6</sup> **Mohsenin N N** Physical properties of plant and animal materials. New-York: Gordon and Breach, 1986
- <sup>7</sup> **Delwiche M J** Theory of fruit firmness sorting by impact forces. Transactions of the ASAE 1987, **30**(4): 1160–1171
- <sup>8</sup> **Brusewitz G H** Drop impact testing applications to fruit quality. Proceedings of the 5th International Conference on physical properties of agricultural materials, Bonn, Germany, 1993
- <sup>9</sup> **Ruiz M** Non-destructive quality measurement and modelling in fruits. Proceedings of the 5th International Conference on physical properties of agricultural materials, Bonn, Germany, 1993
- <sup>10</sup> **Tukey J W** Some graphical and semigraphic displays. Statistical papers in honor of George W. Snedecor. Iowa State University Press: T. A. Bancroft, Ames, 1972
- <sup>11</sup> **Maung K** Measurement of association in a contingency table with special reference to the pigmentation of hair and eye color of Scottish school children. Annals of Eugenics, 1941
- <sup>12</sup> **Goodman L A; Kruskal W H** Measures of association for cross classifications. Journal of American Statistics Association 1954, **49**: 732–764
- <sup>13</sup> **Cohen J** A coefficient of agreement for nominal scalar. Educational and Psychological Measurement 1960, **20**: 37–46
- <sup>14</sup> **Hall L D; Nauda A** An interactive approach for selecting IR&D projects. IEEE Transactions of Engineering Management 1990, **37**(2): 126–133
- <sup>15</sup> **Holmbom P; Pedersen O; Sandell B; Lauber A** Fusing sensor systems: promises and problems. Sensor review, 1989, **9**(3): 143–152
- <sup>16</sup> **Tou J T; Gonzalez R C** Pattern recognition principles. Reading MA: Addison-Wesley Publishing Co., 1974
- <sup>17</sup> **Hall L D** Mathematical techniques in multisensor data fusion. Boston: Artech House, 1992
- <sup>18</sup> **Abdulhafour M; Abidi M A** Data fusion through non-deterministic approaches: a comparison. Proceedings of the SPIE Conference, Sensor fusion VI 1993, **2059**(35): 37–53
- <sup>19</sup> **Anderberg M R** Probability and mathematical statistics. New York, London: Academic Press, 1973
- <sup>20</sup> **Roux M** Algorithmes de classification. Paris: Editions Masson, Collection Méthodes et Programmes, 1985