

High speed intelligent classifier of tomatoes by colour, size and weight

J. Clement¹, N. Novas¹, J. A. Gazquez¹ and F. Manzano-Agugliaro^{2,*}

¹ *Departamento de Arquitectura de Computadores y Electrónica. Universidad de Almería, Almería. Spain*

² *Departamento de Ingeniería Rural. Universidad de Almería, Almería. Spain*

Abstract

At present most horticultural products are classified and marketed according to quality standards, which provide a common language for growers, packers, buyers and consumers. The standardisation of both product and packaging enables greater speed and efficiency in management and marketing. Of all the vegetables grown in greenhouses, tomatoes are predominant in both surface area and tons produced. This paper will present the development and evaluation of a low investment classification system of tomatoes with these objectives: to put it at the service of producing farms and to classify for trading standards. An intelligent classifier of tomatoes has been developed by weight, diameter and colour. This system has optimised the necessary algorithms for data processing in the case of tomatoes, so that productivity is greatly increased, with the use of less expensive and lower performance electronics. The prototype is able to achieve very high speed classification, 12.5 ratings per second, using accessible and low cost commercial equipment for this. It decreases fourfold the manual sorting time and is not sensitive to the variety of tomato classified. This system facilitates the processes of standardisation and quality control, increases the competitiveness of tomato farms and impacts positively on profitability. The automatic classification system described in this work represents a contribution from the economic point of view, as it is profitable for a farm in the short term (less than six months), while the existing systems, can only be used in large trading centers.

Additional key words: artificial vision; low cost; *Solanum lycopersicum*; trading.

Resumen

Clasificadora inteligente de tomates de alta velocidad por color, tamaño y peso

En la actualidad la mayoría de los productos hortofrutícolas son clasificados y comercializados según unos estándares de calidad, que proporcionan un lenguaje común para productores, empacadores, compradores y consumidores. De todas las hortalizas cultivadas en invernadero, el tomate es el que más presencia tiene en superficie y toneladas producidas. En este trabajo se desarrolla y evalúa un sistema de clasificación de tomates de bajo coste con dos objetivos: que sirva a los productores de tomate y que sea capaz de clasificar según los estándares establecidos. En este trabajo se desarrolla un clasificador inteligente de tomates según su peso, diámetro y color. Este sistema optimiza los algoritmos de análisis para el caso específico del tomate, con la ventaja del uso de electrónica de bajo coste. El prototipo es capaz de conseguir la clasificación a velocidad muy elevada, 12.5 clasificaciones por segundo, y utilizando para ello equipos comerciales accesibles y de bajo coste. Se consigue disminuir 4 veces el tiempo de la clasificación manual y no es sensible a la variedad de tomate clasificado. Este sistema facilita los procesos de estandarización y control de calidad, permitiendo aumentar la competitividad de las explotaciones agrícolas de tomate, repercutiendo positivamente en su rentabilidad. El sistema de clasificación automática descrito en este trabajo es una contribución desde el punto de vista económico, ya que se puede amortizar a corto plazo, menos de seis meses, mientras que los sistemas comerciales actuales solo lo hacen en grandes centros de manipulación.

Palabras clave adicionales: bajo coste; clasificación; *Solanum lycopersicum*; visión artificial.

Introduction

The world has 1.6×10^6 ha of greenhouses and macro tunnels (Agugliaro, 2007); the main area is East

Asia with 80% (Marquez *et al.*, 2011), followed by the Mediterranean area with 0.19×10^6 ha (Callejon-Ferre *et al.*, 2009). One of the largest concentrations of greenhouses in the Mediterranean are the coastal areas of

*Corresponding author: fmanzano@ual.es
Received: 28-07-11. Accepted: 27-04-12

Almería, Murcia and Granada, in the Southeast of Spain with approximately 37,500 ha dedicated mainly to intensive vegetable production (Manzano & Cañero, 2010).

The tomato (*Solanum lycopersicum* L.) is the main greenhouse crop by area cultivated and tons produced (Callejon-Ferre *et al.*, 2011). Greenhouse horticultural production in southern Spain, once sorted and packaged, is usually carried from the farm to the marketing center, cooperative or vegetable exchange (auction). The quality assessment of horticultural products is usually carried out by human inspection (Syahrir *et al.*, 2009), which involves a high degree of subjectivity due to psychological factors (Zheng *et al.*, 2006), as there are certain limitations inherent in our visual perception (Gunasekaran & Ding, 1994), visual stress and tiredness. Due to the high growth experienced in electronics, computers, automated systems and machine vision, and the high degree of utilisation of these technologies in society, it is possible to make an intelligent control system that meets all the requirements for classification of fruits and vegetables (Singh *et al.*, 1993), and achieves greater precision and objectivity in the measurement of quality (Alfatni *et al.*, 2011). Today, these techniques are already used in a variety of food industry sectors such as bakery, meat processing, production of ready meals (Sebestyen *et al.*, 2008).

Within the classification of fruits and vegetables, systems for many other types of fruit have also been described: apples (Leemans *et al.*, 1998, 1999; Leemans & Destain, 2004), peaches (Miller & Delwiche, 1989), pears (Hong *et al.*, 2007; Mery & Pedreschi, 2005), different types of citrus (Pencue-Fierro *et al.*, 2004) and bananas (Mendoza & Aguilera, 2004).

The specific case of the automatic classification of the tomato has been studied by several authors. The first to do so were Smith & O'Brien (1979), Butterworth & Butterworth (1979) and O'Brien & Garret (1984) with weight classification systems. The classification of images as an estimate of maturity has been studied by Choi *et al.* (1995), Jahns *et al.* (2001), and Syahrir *et al.* (2009) among other authors while classification by size and weight (for the cherry tomato) was studied by Fernandes *et al.* (2007). All of these were designed for large marketing centers.

The development of pattern recognition algorithms for sorting by size, shape, colour and defects of the surface of tomatoes by digital image analysis of the tomato was started by Sarkar & Wolfe in 1985. There are several studies that provide the image processing

algorithms considered most appropriate for the assessment of tomatoes and do so with great accuracy (Laykin *et al.*, 2002), even recommending what type of sorting algorithm is most efficient and achieves fewest errors (Asadollahi *et al.*, 2009). This paper will present the development and evaluation of a low investment classification system of tomatoes with these objectives: to put it at the service of producing farms and to classify for trading standards.

Material and methods

Selection criteria

To minimise processing time and the amount of information that must be dealt with by an automatic sorter of tomatoes, the minimum number of variables of the fruit considered necessary to make an objective assessment of its quality, and of how precisely each should be sampled has been selected. In the case of tomatoes, these characteristics are size and colour.

Currently there are normalisation tables for the size of tomatoes (OJ, 2009) that are standard in Europe and that differentiate between sizes according to diameter (Table 1). Observance of the sizing scale is compulsory for tomatoes classed "Extra" and I, but this does not apply to trusses of tomatoes. The size tolerance in all categories is: 10% by number or weight of tomatoes conforming to the size immediately below or above the size specified.

As for the colour classification, it is necessary to distinguish between various levels going from green to red, through a series of intermediate green-red levels. These colours are indicative of the ripeness of the fruit depending on variety, and allow the distinction in some cases between tomatoes classed "Extra" or I and lower categories. The European legislation (OJ, 2009) does not explicitly distinguish the types of colour for toma-

Table 1. Size according to diameter

Size code	Denomination	Diameter scale (mm)
10	GGG	> 102
9	GG	> 81 ≤ 102
8	G	> 66 ≤ 81
7	M	> 56 ≤ 66
6	MM	> 47 ≤ 56

Source: OJ (2009).

toes, but USA law does (USDA, 1991), with six different shades of colour.

This paper does not take into account the physical defects in the fruit, as usually the packaging of the production is done manually and the fruits that are defective are discarded directly by the workers.

The criteria of size and colour are considered sufficient to make a broad classification of the tomato production of an agricultural holding. However, it is advisable to also include weight as a criterion for classification, as these variables are taken into account in the following situations:

— As an approximation of the size of tomato in systems without sorting by diameter, due to the low cost of weighing equipment in comparison with industrial vision equipment. It should be noted that there are studies that relate the weight of the tomatoes with their area, calculated using 2D image with an error of 2% (Jahns *et al.*, 2001).

— The weight can also be used as a final estimation of the amount produced in each category, thereby facilitating traceability and control of batches of products.

— There are also markets which apply weight rather than diameter as a classification criterion for distinguishing between sizes.

The values in the normalised chart of sizes (Table 1) are fairly standardised, but unlike these, the criteria for classification by color and weight vary during produc-

tion according to market demand and weather. Therefore, the prototype has been equipped with the functionality of changing any of the three criteria of differentiation (diameter, weight and colour), so that the producer can make a personalised classification of the tomato according to diameter and colour, or according to weight and colour.

Global structure of the system

To help explain how the system works, its overall structure has been divided into several subsystems: mechanical elements of transport and support; weighing subsystem; vision subsystem; synchronisation and ejection central subsystem; and monitoring subsystem.

Except for the subsystem of transport and support, which is purely mechanical, all the electronics of the machine are included in the others and within each there is an element dedicated to processing data and information. Thus, we have opted for distributed control architecture, allowing the use of lower performance controllers, and these communicate with each other using standard protocols such as Ethernet or RS485, whose speed is sufficient, and which are widespread in industrial control systems. Also when using multiple computing cores, multiple tasks can be carried out in parallel, so that the total processing time decreases. In Fig. 1 we

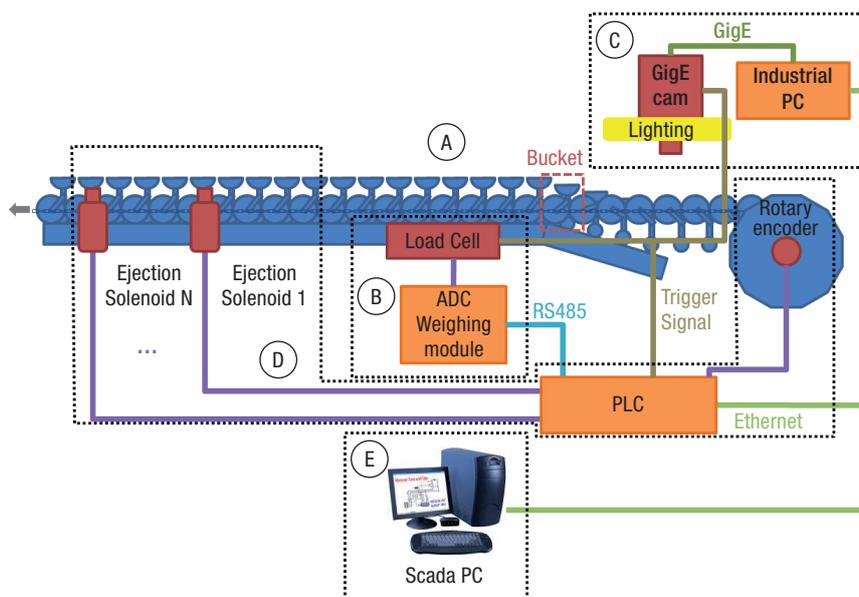


Figure 1. General block diagram. Subsystems: A) transportation and support, B) weighing, C) vision, D) control of synchronisation and ejection, E) monitoring.

have presented a general block diagram of the prototype developed.

Mechanical elements for transport and support

The system has been developed on a mechanical chassis on which the product to be classified circulates in buckets deposited one by one. These are subjected to a mechanised chain that moves along the entire structure of the machine so that the product can be pushed out of the bucket at any point along its route.

The structure provides the support necessary to place both sensors (cameras, encoders and load cells) and actuators (ejection solenoids, lighting and motors).

It is important to highlight the importance of the design of the bucket. Fig. 2 shows a three-dimensional representation of it to enable a better understanding of its functionality and advantages. This design makes it possible to tilt the bucket for the expulsion of tomatoes, so care is taken to avoid damage by this system, but it also has two other very important functions. The first is to raise the tomatoes onto a platform with a single point of support on which the weight of the tomato-support assembly rests, allowing the weighing process to be performed as it passes through the load cell. The second function is to enable the rotation of the tomato in order to take sample pictures around it from a fixed camera position. This is possible by lowering the platform until it fits into the cylindrical base, placing the tomatoes between two contiguous cylinders that are continuously rotating, causing the tomatoes to rotate too.

Weighing subsystem

The weighing subsystems used in fruit sorting machines usually use load cells as sensors (Frances *et al.*, 2000). In this case we used a load cell manufactured by Utilcell (2011), specifically the 260 model, which consists of a double bending beam load cell with viscous damping, of 5 kg capacity, accurate to 4,000 divisions OIML R60 class C (O.I.M.L. R60, 2000), made of stainless steel and with a degree of IP66 protection, as this has the advantages of allowing rapid stabilisation, high speed weighing, vibration reduction and an extension of the life of the cell. This cell makes less than 2-g of error in the weight measurement, which is considered acceptable for the classification of medium-sized fruits and vegetables like tomatoes.

In the type of weighing scheme used (see Fig. 3), a platform decoupled from the rest of the mechanised chain or conveyor belt is used, which is very common in the industry. The length of this dynamic weighing platform can be as long as the separation distance between two adjacent buckets, which in this case is 19 cm. The platform needs to be as long as possible, since in this way the time over that platform is maximised and measurement will be more stable.

In order to find out the level of processing required by the signal from the load cell, and give an objective value of the weight, it is necessary to know the waveform of the signal through a typical weighing cycle. Using this type of load cell and sampling sufficiently fast, the Fig. 4 shows a typical weighing cycle. In this graph, overshoots occurring due to the entrance and exit of the object on the weighing platform, destabilising the measurement, can be observed.

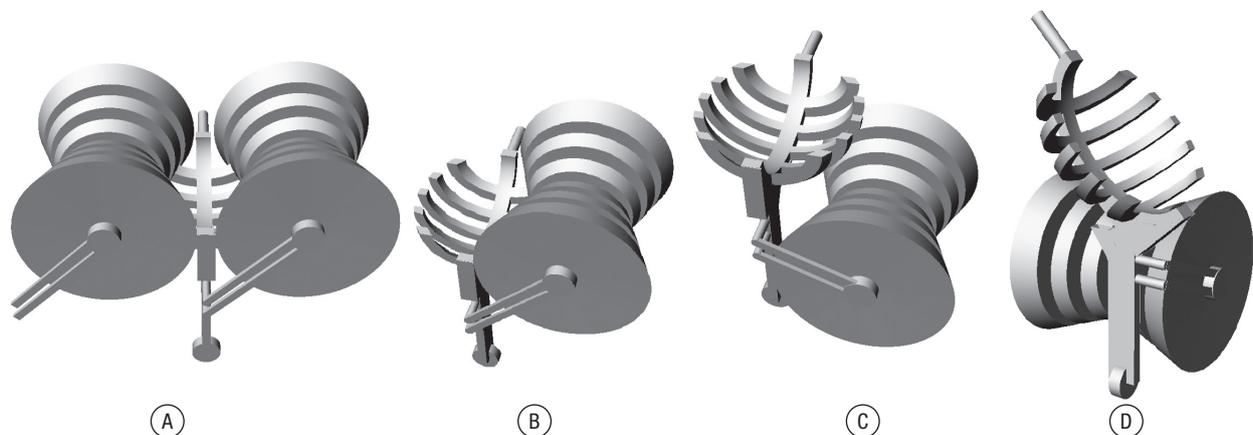


Figure 2. Detailed three-dimensional bucket. Positions: A) two adjacent buckets fitting (rotation position for image sampling), B) view of a bucket (rotation position for image sampling), C) weighing, D) tilt (expulsion of fruit).

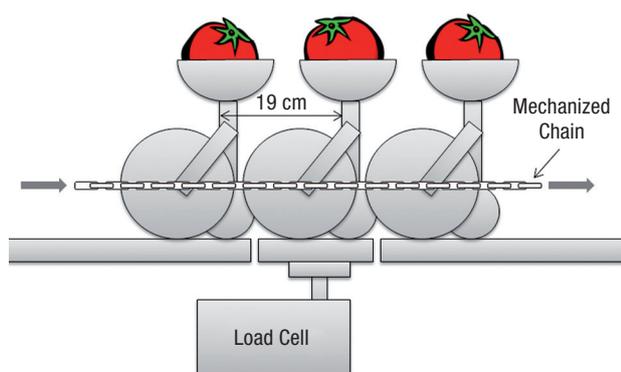


Figure 3. Scheme of dynamic weighing subsystem.

To prevent these disturbances interfering with the measurement, we could use a low-pass filtering of the signal, but this solution has the disadvantage of introducing an additional delay in the measurement. The solution adopted is to perform an oversampling of the signal from the load cell to calculate the average value, starting this measurement when the bucket reaches a certain point on the weighing platform, which coincides with the instant when overshoots are considered over, and which will last until the bucket's exit from the platform. The synchronisation system is responsible for calculating the exact position of the bucket at any time and therefore this will trigger the signal to begin measuring.

The acquisition system should be able to perform a correct sampling of the information in order to ensure the validity of the information that comes from the weight sensor. A commercial weighing module made by Hauch & Bach (2011) was used, namely the 78.1 LDU model, which is specifically designed for high speeds,

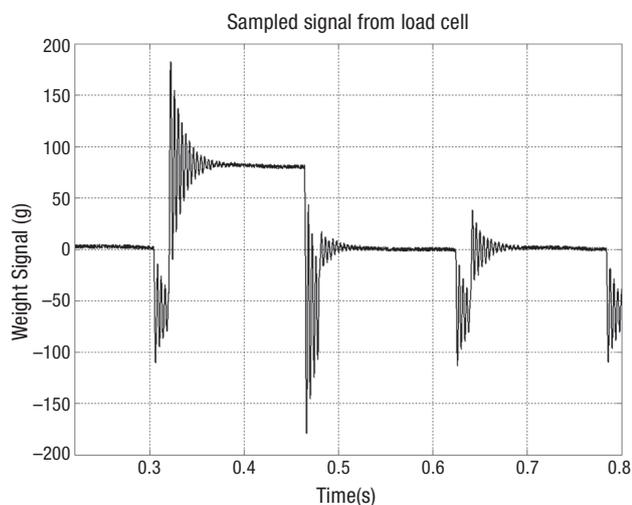


Figure 4. Typical weighing cycle.

because it has a resolution of 18 bits and is capable of 2400 conversions per second. It also has a RS422/485 communication port that makes it easy to integrate it into any network control, and has the functions for signal processing necessary for our application (filtering, averaging, etc).

Vision subsystem

This part of the machine consists of an industrial PC that performs the processing of images captured with a camera interface GigE. In addition the PC will communicate via Ethernet with the rest of the control system, sending the calculated value of the diameter and colour of tomato on request.

To capture images, a colour CCD camera from the manufacturer Dalsa (Genie C640) has been selected. This has a global shuttering system specifically designed to capture moving objects, minimising image blur. The camera can capture up to 64 frames s^{-1} with a resolution of 640×480 pixels, which is more than enough for the imaging system. This model allows both the aperture of the lens and the exposure time to be set, thereby enabling very low capture times to be achieved. In this case it has been programmed to work with exposure times of 12 ms. It also has a digital input that can be used as an external trigger, which is sent by the synchronisation system.

Lighting is also an important part of any imaging system, because a good lighting system can facilitate the subsequent processing required (Novini, 1993). In the case of tomatoes, they should be illuminated with white light to keep the tonality of the colour, and do so in a way that produces no glare, as this can complicate the processing. A LED lighting system that incorporates a diffuser to eliminate glare on reflective surfaces such as tomatoes has been used. The camera has been placed with two lights on the sides, within a closed compartment that isolates it from external light. For the whole of its trajectory through the interior of the compartment, the bucket will be in position B (see Fig. 2), which in turn will cause the rotation of the tomato, so that several screenshots can be made around each fruit to obtain samples of its entire girth.

Moltó *et al.* (1998) found that with 4 images of apples, up to 80% of its surface could be observed, although the classification rate was 0.25 fruits s^{-1} . It was decided to capture in a single image up to four consecutive buckets, repeating this process every time a

new bucket entered the bay. Thus, a tomato's image is captured four times, from when it goes into the bay until it comes out (see Fig. 5).

In order to obtain the greatest possible contrast, black has been used for the background image, highlighting objects of interest in the background. Therefore, the buckets are made of black material and the rest of the background visible from the camera is covered with matt black materials or paint. To simplify the representation and description of the objects of interest, tomatoes have been modelled as ellipses whose interior colour takes different shades from green to red. Thus, to approximate its size, it is sufficient to obtain the maximum and minimum diameters of the ellipse that is represented, and to describe the colour with just a numeric value, where zero corresponds to the colour red and one hundred to green. The algorithms of image processing were performed using the open source computer vision library OpenCV, as this has many functions that cover the system requirements (see Fig. 6).

Diameter calculation

The shape of agricultural products such as fruit, vegetables and grain is one of the most important factors for their classification and grading with regard to commercial quality (Morimoto *et al.*, 2000). Costa *et al.* (2011) analysed the state of the art of the agricultural product shape analysis for computer vision implementation and concluded that the tomato was classified according to 10 shape categories such as rounded, high-rounded, ellipsoid or pyriform.

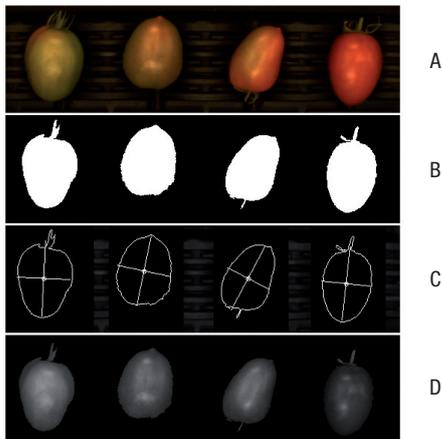


Figure 5. Image processing on industrial PC. A) colour image, B) threshold, C) edge, centre and diameter detection, D) green channel masked.

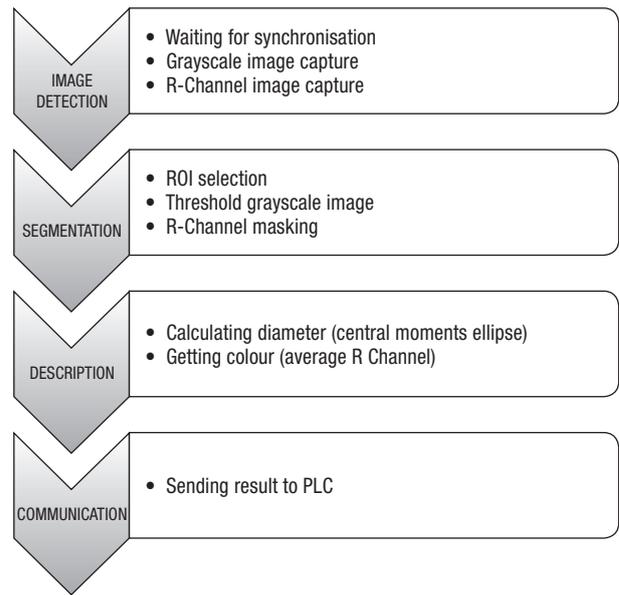


Figure 6. Duty cycle of image processing PC.

In this case it was decided to use the ellipsoidal shape. To calculate diameter, we used the grayscale image where the four regions of interest of the image corresponding to four tomatoes inside the bay are identified. These four regions have a fixed location and size, since capture is performed based on a sync signal that is triggered when the buckets reach a certain position.

Having identified the regions of interest, a threshold process was carried out for each one, so that the background could be removed, keeping only the product shape, in order to subsequently perform edge detection. Once we had the outlines of each piece, they were compared with an ellipsoid, so that both maximum (D) and minimum (d) diameter as its centre (Cx, Cy) can be approximated using the central moments, m_{ij} (Rocha *et al.*, 2004):

$$C_x = \frac{m_{10}}{m_{00}} \quad [1]$$

$$C_y = \frac{m_{01}}{m_{00}} \quad [2]$$

$$D = \sqrt{\frac{2 \cdot \left(m_{20} + m_{02} + \sqrt{(m_{20} - m_{02})^2 + 4 \cdot m_{11}^2} \right)}{m_{00}}} \quad [3]$$

$$d = \sqrt{\frac{2 \cdot \left(m_{20} + m_{02} - \sqrt{(m_{20} - m_{02})^2 + 4 \cdot m_{11}^2} \right)}{m_{00}}} \quad [4]$$

Colour calculation

In the case of tomatoes, colour can vary from shades of green to red. In an RGB colour image, the level of green can be approximated using only the value of the pixels of the green channel (G). To avoid the inclusion of unwanted pixels of the image, the threshold performed in the diameter calculation is used to mask the background. So there will be a grayscale image whose pixels indicate the level of green in the image except for the background, whose pixels have a null value. Calculating the average level of grey in each region of interest can approximate the colour of each of the tomatoes.

As final value for the colour, the average values calculated from the four images of the fruit from when it enters the bay until it comes out have been selected, while for the diameter it will remain the highest value obtained in the four measures.

Synchronisation and ejection central subsystem

A modular PLC (Q series) from the manufacturer Mitsubishi (2011) has been selected as the core of this subsystem, for several reasons: i) it has a very high processing speed: 9.5 ns/logical instruction; ii) being a modular system, it could extend its operation to control up to 4096 input points and 8192 outputs, more than enough to control motors and ejectors; iii) it also has a dedicated interface for motion control and high-speed sync; iv) it has communication interfaces RS485 and Ethernet.

The expulsion of the products is accomplished by a solenoid that will force the tilting of the bucket with a resulting tomato drop, and will be activated by the controller at the right time.

In order to obtain any sync signal, it is necessary to use a sensor to identify the movement of the buckets. In this case we used an incremental and programmable rotary encoder, from the manufacturer Sick (2011). This sensor was placed on the movement shaft of the main engine of the machine and was connected to the sync module of the PLC. The programmable feature of this model allows the number of pulses corresponding to 360° to be defined. This will make it much easier to adapt to any size of dish of the mechanised chain.

The core of this subsystem should undertake the task of synchronising the processes of image capture, weigh-

ing and activation of ejectors. To do this we have defined two sync signals. The first (S1) indicates the time of opening and closing the expulsion mechanisms. The second (S2) indicates the beginning of the process of weighing and image capture and processing. Fig. 7 shows a timing diagram which represents the two signals.

The measurement and characterisation processes that take place in weighing and vision subsystems will be released at once. These processes require a computation time that must not be interrupted by requests from the central system. Therefore communication between subsystems to exchange data, either via RS485 or Ethernet, is performed in each cycle but always after the process is complete. At this point weight, diameter and colour data will be ready for transmission. Using alternate timing signals S1 and S2 (see Fig. 7) optimises the use of the controllers, resulting in greater speed and a lower error rate in data transactions. The physical placement of load cells, cameras and ejectors directly influences the activation time of the sync signals. Therefore the relative position of these elements plays an important role, since it must allow the maximum distance between S1 and S2 pulses in order to minimise the total duration of a cycle.

The duty cycle of the central control subsystem for synchronisation and expulsion is divided into three main tasks:

- Positioning and synchronisation. This is the main task, which generates the timing signals and performs the monitoring and recording of the current position of the machine.
- Operation. General management of the external controls of the machine and engines (start, stop, emergency stop, etc.).
- Classification. This task is released at the time indicated by task a), as it depends on the condition of synchronism of the machine.

The classification task is the most complex (see Fig. 8). The products placed in buckets circulating

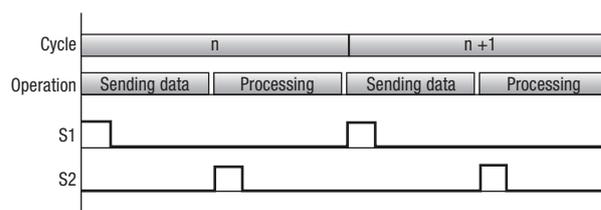


Figure 7. Diagram of sync. Operation: time distribution in weighing and vision systems. S1: expulsion trigger. S2: image capture and weight trigger.

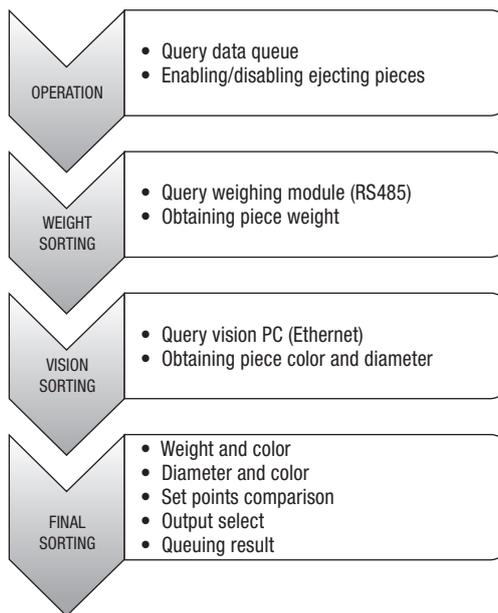


Figure 8. Disaggregation of the classification task.

through the machine, which are analysed as they pass through the load cells and cameras, must be expelled. First, this action is performed with the activation or deactivation of the ejectors according to the classification carried out previously. This is followed by the corresponding requests for weight data via RS485, and diameter and colour via Ethernet. Having obtained these data, the system is ready to classify the product just out of the subsystems of measurement, which will be assigned a particular class that will correspond with an output for expulsion. This value will be stored in a queue to be consulted at the start of the corresponding cycle and will activate their corresponding ejector.

The classification of a fruit is based on two criteria agreed a priori. The options are diameter and colour sorting, or weight and colour sorting. The use of weight and diameter simultaneously as classification criteria can lead to inconsistencies and therefore is not considered an option. In any case, the weight data is stored and transmitted to the monitoring system to help control product traceability.

Monitoring subsystem

This subsystem is based on a custom-designed Scada in C++. Through this we have total control in real time of the central control subsystem for sync and expulsion, and therefore of any actuator or sensor of

the machine. It uses an Ethernet network to communicate with the PLC.

The Scada design allows the overall control and supervision of the system, but can also store different work schedules, edit and assign. The classification criteria and set points with which you want the machine to run are stored in such programmes.

Another added functionality is to store information on the number of products, average weight, average diameter, average colour, total weight, and so on. This information can be viewed as a whole, according to classification criteria, and can also be printed on a dispatch note with configurable format.

Evaluation methodology

To assess the correct operation of a tomato sorting machine, some authors use 50 tomatoes for the colour (Syahrir *et al.*, 2009). This paper has proposed using 50 kg of tomatoes. M caliber usually weighs about 80 g, thus 10 times more tomatoes are used for evaluation.

It was decided to perform different tests by dividing a quantity of tomatoes into various classes based on criteria of determined colour, weight and diameter. This evaluation took place in a real working environment, a farm in Almeria (southern Spain). Three different kinds of tomatoes were used: LongLife (Daniela variety), plum tomatoes (Mirade variety) and green tomatoes (Cecilio variety). Each of these varieties has different market requirements and classification criteria, which allows the machine to evaluate several scenarios.

For each variety, errors based on the rating criterion selected were evaluated. For Daniela and Mirade, varieties that are harvested when the fruit is orange or red, the product was classified into three shades by colour (green, orange, red), while for the Cecilio variety, as this is harvested green, you only had to distinguish red tomatoes from green, so it was sorted by colour into two shades (green and red). For all three varieties, the diameter and weight sorting was divided into five categories according to Table 1 (MM, M, G, GG, GGG). The weights for each product were selected so that they corresponded to these sizes.

The productivity assessment was carried out by evaluating the total time it took to classify 1000 kg of each of the previous three tomato varieties at full speed, and comparing the result with the amount of time it took to classify these tomatoes by hand with four expert workers.

Results

Classification

Table 2 shows the results of the classification tests. Accuracy is presented by depicting the correct classification percentages, as well as the percentage of errors when classifying to adjacent classes at one or more than one level of proximity (several classes of distance).

Percentages indicate that classification by colour only produces errors of one adjacent class apart. In the case of the Cecilio variety, although the product is divided into fewer classes, the increase in errors is due to the homogeneity of the colour being less than in other varieties.

The classification by diameter is very accurate and produces few mistakes. There were some failures because at full speed, some products were not properly placed in the buckets.

In the case of sizing according to the weight of the fruit, errors are more numerous than with sizing according to diameter. This is because of interference caused at the point of input and output of the product on the weighing platform, along with the mechanical vibrations of the machine, which are accentuated at higher speed. Also it should be mentioned that when driving the machine at such high speed, there is a continuous movement of the fruit in the bucket, which also disturbs the measurement of weight, but less so the colour and diameter.

To check the maximum speed of operation, it was increased progressively while checking that the error rate did not increase. The critical point is reached when the cycle time has been shortened enough for the processing and sending times to overlap. In this situa-

tion the system is not able to respond quickly enough to tilt the buckets. With proper synchronisation settings, the system was optimised to reach a speed of 12.5 classifications per second and per line (see Fig. 9). At speeds above this, the system starts to cause errors with products that remain unclassified.

Productivity

The results of the comparison of productivity are summarised in Table 3. It should be noted that the yields are positioned above 3000 kg h⁻¹, 3 to 5 times higher than in the classification manual. Also there is no appreciable difference in time between the automatic classifications of different varieties of tomato, while as expected, there is in manual sorting.

One can approximate the number of people needed to match the output of the prototype, since we are comparing the result of automatic classification to the time it takes four expert workers, which would be 16 people for the Daniela variety, 14 for Miriade and 21 for Cecilio. Studying the previous data of operation time, comparing the automatic and manual sorting and taking into account the average cost of workers' wages (assuming a net monthly salary of € 1,000), with a system cost of € 50,000, shows that in the best case for the Cecilio variety, the money would be recouped in 2.4 months, while in the worst case for the Miriade variety, it would be recouped in 3.6 months.

Discussion

Systems that automate tasks traditionally performed manually always represent an advance, both economi-

Table 2. Misclassification according to varieties and criteria

Variety	Test	Exact (%)	± 1 Missed (%) ^a	Severe missed (%)
Daniela	Color	96	4	0
	Diameter	97	< 3	< 1
	Weight	83	15	2
Miriade	Color	97	3	0
	Diameter	95	5	0
	Weight	85	13	2
Cecilio	Color	93	7	0
	Diameter	96	4	0
	Weight	80	17	3

^a ± 1 Missed: misclassification by one adjacent class.

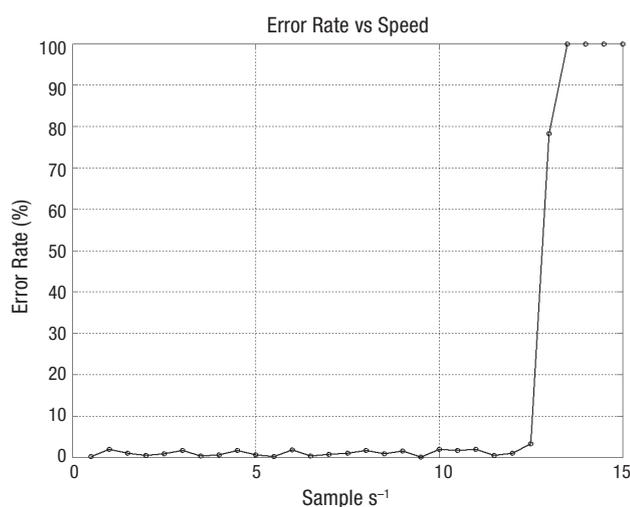


Figure 9. Error rate depending on the operating speed of the system.

cally and socially. In this case the automatic classification system described in this work represents a contribution from the economic point of view, as it is profitable for a farm in the short term (less than six months), while the existing systems, such as those developed by other authors (Smith & O'Brien, 1979; Butterworth & Butterworth, 1979; O'Brien & Garret, 1984; Choi *et al.*, 1995; Jahns *et al.*, 2001; Syahrir *et al.*, 2009; Fernandes *et al.*, 2007) can only be used in large trading centers.

Current commercial systems such as Greefa are also thought for large trading centers because of the price. Our system is five times cheaper. This makes it possible to install it on farms of average size of 2.5 ha that is very common in southeastern Spain (Marquez *et al.*, 2011).

On the other hand, a very high speed of operation has been achieved with commercial modules, which means very high levels of productivity compared to traditional or manual systems. These levels could be multiplied if several parallel lines of buckets were used in the same mechanic chassis.

It has been shown that the levels of accuracy of machine vision systems for such applications of veg-

etable classification are sufficiently good to promote their use. In addition, the technology is mature enough to achieve satisfactory results with a low investment.

The prototype developed is flexible, allowing the automated system to adapt quickly to the sorting parameters in order to suit market requirements.

The use of weighing systems in this system for classification by size returns results that are not as favourable as computer vision, but still acceptable in most cases. Also, weighing is an economical method of approximating size, so its use is only defended in very low cost classification systems, when vision systems are not used, this being justified by the close to 100% correlation between the weight and the 2D image of the tomato (Jahns *et al.*, 2001).

In the manufacture of this prototype, simple algorithms have been chosen, that maximise the processing speed by reducing the complexity of operations performed, while maintaining the error rate and quality parameters at an acceptable level, thus achieving higher productivity. Moreover, by reducing the complexity of the algorithms, the process can be performed by the commercial electronic equipment available, and which has fewer requirements than those designed with systems similar to the ones considered by Laykin *et al.* (2002) or Asadollahi *et al.* (2009) among others.

Some of the systems referenced could classify 0.25 parts s⁻¹ (Moltó *et al.*, 1998), while this system has optimised the necessary algorithms for data processing in the case of tomatoes, so that productivity is greatly increased, with the use of less expensive and lower performance electronics. Modern equipment can sort up to 10 fruits s⁻¹, differentiating color, size, shape (Lino *et al.*, 2008), our system achieves the same speed and also taking into account the weight.

As a final conclusion it should be noted that this system will have a clear impact on the quality of the final classification of the marketed product, since the subjectivity of classification by human inspection is completely avoided (Gunasekaran & Ding, 1994; Zheng *et al.*, 2006; Syahrir *et al.*, 2009). It also improves the

Table 3. Comparison of productivity. T: time to process 1,000 kg of product. P: productivity

Tomato	T ₁ Manual (min)	T ₂ Auto (min)	P ₁ Manual (kg h ⁻¹)	P ₂ Auto (kg h ⁻¹)	P ₁₂ Improvement (%)
Daniela	82	20	731	3,000	410
Miriade	76	22	789	2,727	345
Cecilio	93	18	645	3,333	516

profitability of the farm, as the amortization of the equipment is short-term (Meyers, 1988).

Acknowledgements

This research was partially funded by the project RAMAS from Research and Technology National Plan, Funds No. TEC2009-13763-C02-02 and supported by the Research Group TIC-019 “Electronic, communications and Telemedicine” from University of Almería (Spain). We would like to thank the company I. A. Madalba S. L. for their contribution to the system’s mechanical design and construction. We also would like to thank Barbara Lamplugh for her help with the translation.

References

- Agugliaro FM, 2007. The gasification of vegetable residues of greenhouses for making electrical energy in the south of Spain: his ubication study by GIS. *Interciencia* 32(2): 131-136.
- Alfatni MS, Shariff ARM, Abdullah MZ, Ben Saeed OM, Ceesay OM, 2011. Recent methods and techniques of external grading systems for agricultural crops quality inspection — Review. *Int J Food Eng* 7(3): art. no. 10. doi:10.2202/1556-3758.1932.
- Asadollahi H, Kamarposhty MS, Teymoori MM, 2009. Classification and evaluation of tomato images using several classifier. *Computer Science and Information Technology - Spring Conf IACSITSC '09*, pp. 471-474.
- Butterworth HM, Butterworth WR, 1979. Electronic weighing and recording system for tomato grading. *ASAE-CSAE Summer Meeting*, Winnipeg, Manitoba, Canada. June 24-27, pp: 4-8.
- Callejon-Ferre AJ, Manzano-Agugliario F, Diaz-Perez M, Carreño-Ortega A, Perez-Alonso J, 2009. Effect of shading with aluminised screens on fruit production and quality in tomato (*Solanum lycopersicum* L.) under greenhouse conditions. *Span J Agric Res* 7(1): 41-49.
- Callejon-Ferre AJ, Velazquez-Marti B, Lopez-Martinez JA, Manzano-Agugliario F, 2011. Greenhouse crop residues: Energy potential and models for the prediction of their higher heating value. *Renew Sust Energ Rev* 15(2): 948-955.
- Choi K, Lee G, Han YJ, Bunn JM, 1995. Tomato maturity evaluation using colour image analysis. *T Am Soc Agric Eng* 38(1): 171-176.
- Costa C, Antonucci F, Pallotino F, Aguzzi J, Sun D-W, Menesatti P, 2011. Shape analysis of agricultural products: a review of recent research advances and potential application to computer vision. *Food Bioprocess Technol* 4(5): 673-692.
- Fernandes C, Corá JE, Braz LT, 2007. Grading of cherry tomatoes based on fruit size and weight. *Horticultura Brasileira* 25(2), 275-278.
- Frances JV, Calpe J, Martinez M, Rosado A, Serrano, AJ, Calleja J, Diaz M, 2000. Application of ARMA modeling to the improvement of weight estimations in fruit sorting and grading machinery Acoustics, Speech, and Signal Processing. *Proc 2000 IEEE Int Conf*, Vol 6. pp: 3666-3669.
- Gunasekaran S, Ding K, 1994. Using computer vision for food quality evaluation. *Food Technol* 6: 151-154.
- Hauch and Bach, 2011. Load cell digitizing unit LDU78.1. Available online in <http://www.haubac.com/haubac.asp?p1=54> [15 February 2011].
- Hong T, Qiao J, Ning W, Ngadi MO, Zhao Z, Li Z, 2007. Non-destructive inspection of Chinese pear quality based on hyperspectral imaging technique. *T Chinese Soc Agric Eng* 23(2): 151-155.
- Jahns G, Moller-Nielsen H, Wolfgang P, 2001. Measuring image analysis attributes and modelling fuzzy consumer aspects for tomato quality grading. *Comput Electron Agr* 31(1), 17-29.
- Laykin S, Alchanatis V, Fallik E, Edan Y, 2002. Image processing algorithms for tomato classification. *T ASAE* 45(3): 851-858.
- Leemans V, Destain MF, 2004. A real-time grading method of apples based on features extracted from defects. *J Food Engg* 61: 83-89.
- Leemans V, Magein H, Destain MF, 1998. Defects segmentation on ‘golden delicious’ apples by using colour machine vision. *Comput Electron Agr* 20: 117-130.
- Leemans V, Magein H, Destain MF, 1999. Defect segmentation on ‘jonagold’ apples using colour vision and a bayesian classification method. *Comput Electron Agr* 23: 43-53.
- Lino ACL, Sanches J, Dal Fabbro IM, 2008. Image processing techniques for lemons and tomatoes classification. *Bragantia* 67(3): 785-789.
- Manzano-Agugliario F, Cañero-Leon R, 2010. Economics and environmental analysis of Mediterranean greenhouse crops. *Afr J Agric Res* 5(22), 3009-3016.
- Marquez AL, Baños R, Gil C, Montoya MG, Manzano-Agugliario F, Montoya FG, 2011. Multi-objective crop planning using pareto-based evolutionary algorithms. *Agr Econ* 42(6): 649-656.
- Mendoza F, Aguilera JM, 2004. Application of image analysis for classification of ripening bananas. *J Food Sci* 69(9): 471-477.
- Mery D, Pedreschi F, 2005. Segmentation of colour food images using a robust algorithm. *J Food Eng* 66: 353-360.
- Meyers JB, 1988 Improving dynamic visual inspection performance, M.Sc. thesis, Univ. of Georgia, Athens.

- Miller BK, Delwiche MJ, 1989. A color vision system for peach grading. *T ASAE* 34(6): 2509-2515.
- Mitsubishi, 2011. PLCs Modulares — Control PLC. Q Series. Available online in <http://www.mitsubishi-automation.es> [15 February 2011].
- Moltó E, Blasco J, Benlloch JV, 1998. Computer vision for automatic inspection of agricultural produces. *SPIE Symp on Precision Agriculture and Biological Quality*. 1-6 November, Boston, MA, USA.
- Morimoto T, Takeuchi T, Miyata H, Hashimoto Y, 2000. Pattern recognition of fruit shape based on the concept of chaos and neural networks. *Comput Electron Agr* 26: 171-186.
- Novini A, 1993. Fundamentals of machine vision lighting. *Proc WESCON '93*. San Francisco, CA, USA, Sept 28-30. pp 44-52. Doi 10.1109/WESCON.1993.488407.
- OIML, 2000. Recommendation of the International Organisation of Legal Metrology, R 60 Metrological regulation for load cells. Edition: 2000.
- O'Brien M, Garret RE, 1984. Development of a mechanical and electronic assisted tomato grading station. *T Am Soc Agr Eng* 27(5): 1614-1618.
- OJ, 2009. Establishing the standard import values for determining the entry price of certain fruit and vegetables. Commission Regulation (EC) No 771/2009 of 25 August 2009 amending Regulation (EC) No 1580/2007.
- Pencue-Fierro L, Patiño-Velasco M, Zambrano-Velasco C, Leon-Tellez J, 2004. Digital image analysis in the design and implementation of a visual inspection system for fruit grading. *Proc of SPIE - The International Society for Optical Engineering* 5622 (Part 1), art. no. 17, pp. 91-96.
- Rocha L, Velho L, Carvalho PCP, 2004. Motion reconstruction using moments analysis. *Proc 17th Brazilian Symp on Computer Graphics and Image Processing*. pp: 354-361.
- Sarkar N, Wolfe RR, 1985. Feature extraction techniques for sorting tomatoes by computer vision. *T ASAE* 28(3): 970-974.
- Sebestyen G, Nicoara D, Rusu M, Buzdugan L, Cerghizan M, Saplacan G, 2008. supervisory control and quality assurance in food industry. *AQTR 2008, IEEE Int Conf on Automation, Quality and Testing, Robotics*, vol 1, pp: 195-199.
- Sick, 2011. Sensor intelligence. Rotary incremental encoders. Available online in http://www.sick.com/group/en/home/products/product_portfolio/encoders/pages/rotary_incremental.aspx. [15 February 2011].
- Singh N, Delwiche MJ, Johnson RS, 1993. Image analysis methods for real-time color grading of stonefruit. *Comput Electron Agr* 9(1): 71-84.
- Smith NE, O'Brien M, 1979. Electronic weighing and recording system for tomato grading. *ASAE-CSAE Summer Meeting*. Winnipeg, Manitoba, Canada, June 24-27, pp: 20-40.
- Syahrir WM, Suryanti A, Connsynn C, 2009. Color grading in tomato maturity estimator using image processing technique. *Proc 2nd IEEE Int Conf on Computer Science and Information Technology*, art. no. 5234497, pp: 276-280.
- USDA, 1991. United States standards for grades of fresh tomatoes. *USDA Agricultural Marketing Service*, Washington, DC.
- Utilcell, 2011. Células de carga monocélula. Modelo 260. Available online in http://www.utilcell.es/secciones/productos/ficha.aspx?seccio=productos&vista=0&categoria=0&tipus=0&model=modelo_260. [15 February 2011].
- Zheng C, Sun D-W, Zheng L. 2006. Recent developments and applications of image features for food quality evaluation and inspection - a review. *Trends in Food Science and Technology* 17 (12): 642-655.