Computer Vision Applied to Flower, Fruit and Vegetable Processing
Luis Gracia, Carlos Perez-Vidal, Carlos Gracia

Abstract—This paper presents the theoretical background and the real implementation of an automated computer system to introduce machine vision in flower, fruit and vegetable processing for recollection, cutting, packaging, classification, or fumigation tasks. The considerations and implementation issues presented in this work can be applied to a wide range of varieties of flowers, fruits and vegetables, although some of them are especially relevant due to the great amount of units that are manipulated and processed each year over the world. The computer vision algorithms developed in this work are shown in detail, and can be easily extended to other applications. A special attention is given to the electromagnetic compatibility in order to avoid noisy images. Furthermore, real experimentation has been carried out in order to validate the developed application. In particular, the tests show that the method has good robustness and high success percentage in the object characterization.

Keywords—Image processing, Vision system, Automation

I. INTRODUCTION

APPROXIMATELY, 900 million tons of vegetables and 500 million tons of fruits were produced in 2005 all over the world. The vegetable production is persistently increasing and the same trade is observed for fruit production. Fruit production grown about an 8% yearly in 2006, 2007 and 2008; this value is quite smaller for vegetables, around 3-4% for the same period. The European Union, Northern America and Japan are the major fruit and vegetable importers, mainly from South America, Russia and Asian countries [1]. At the same time, the flower industry is blossoming all over the world. Despite a sluggish global economy, estimates are that the sales of floral products increased significantly in the first years of the new century. According to Dr. Terri W. Starman (Department of Horticultural Sciences at Texas A&M University), in 2000 the wholesale value of the global flower market came to US$24.68 billion, which translated to a retail value of US$77 billion. Some types of flowers/fruits/vegetables (see Fig. 1) are especially relevant due to the great amount of units that are manipulated and processed each year over the world [2]. In fact, there are many works related with citrus [3] or other more unusual fruits like pomegranate [4].

Beside this, computer or machine vision is becoming more and more used in industry, including agriculture, although in this field it is not enough widespread despite that the world agriculture production is increasing every. This paper presents the theoretical background and real implementation of an automated computer system to introduce computer vision [5] in flower, fruit and vegetable processing for recollection, cutting, packaging, etc. Some basic issues related to image processing are presented by [6] in order to increase machine vision in agriculture automation [7]. This research tackles some technical aspects that are important for automated machine vision [8], such as the camera, the acquisition card, the illumination, the electromagnetic compatibility (filtering and electric connections), etc.

The final target of the developed application is to obtain the position of the flower/fruit/vegetable centre (eliminating the background, peduncle and other non-relevant parts like sheets) and for example to compute the cutting point that allows to separate the element from its peduncle. This information (centre of the element) can be used for several types of agricultural tasks, such as recollection, cutting,
The developed application is tested with different types of flowers, fruits and vegetables to check the effectiveness and robustness of the solution. An important feature of the proposed algorithm is that it is able to work with several elements inside the camera field of view. This adds flexibility to the application in order to work in the country or in a greenhouse, where the elements are very close to each other and the location of all them must be obtained in real time.

The specific objectives of this work are: (1) to describe issues related to computer vision for flower/fruit/vegetable processing; (2) to present the electromagnetic compatibility (EMC) considerations to take into account; (3) to show the final application developed and.

II. CONFIGURATION SETUP

Fig. 2 presents the configuration setup used in this work. The camera, which is inside a Faraday jail, acquires the image of the real flowers/fruits/vegetables. The video signal is sent from the camera to the industrial computer, which is equipped with a frame grabber that captures the image. The computer obtains the position of the centre of the detected objects and this information is used to process the element. Using this object characterization, the computer sends a “Motion command” (depending on the desired task: recollection, cutting, packaging, classification, etc.) through a USB (universal serial bus) link to the “Driver” that controls the “Electric motor” devoted to modify or manipulate the process.

The configuration setup shown in Fig. 2 has a significant problem of EMI due to the proximity of the “source” (electric drive and motor) and “victim” (vision system) of electrical noise. The effects of conducted and radiated interference are represented in Fig. 2. Therefore, all the issues described in next section “Electromagnetic compatibility” have been considered to obtain a proper performance of all the elements in the system. The electromagnetic compatibility directives that are applicable to this work and to similar ones are: EN 50081-1 (residential & light industry) and EN 50081-2 (heavy industry) for Emission Standards and; EN 50082-1 (residential & light industry) and EN 50082-2 (heavy industry) for Immunity Standards and are summarized in next section.

A. Computer Features

**Hardware:** Industrial computer (PC type) for rack mounting; rack type keyboard and mouse; TFT color Display of 15" XGA (1024 x 768); Pentium®4 1,7GHz; 1GB of RAM; 4 serial ports RS-232; 3 USB ports; Hard Disk of 250GB; 3 slot for PCI.

**Software:** Microsoft Windows XP Professional with Service Pack 3; Matrox Image Library (MIL) 8.0; Microsoft Visual C++ 6.0 (programming environment); and Deep Freeze Standard 6.52 software (by Faronics®) to protect the core operating system and configuration files restoring the computer back to its original configuration each time it restarts.

B. Vision System Features

**Camera** (SONY XC-HR50): Monochrome; 1/3 type progressive scan (better than interlaced scans for capturing moving objects) CCD with square pixels; VGA resolution (704x495 pixels) capturing at a speed of 60 full frames/sec.; minimum illumination of 1.0 lux at F1.4; signal/noise ratio of 58 dB; electronic shutter from 1/100 to 1/30,000 sec.

**Lens:** Special for machine vision with C-mount; focal length of 12 mm; maximum aperture ratio of 1:1.4; iris range from 1.4 to 16.

**Frame grabber** (Matrox Meteor II/MC): captures from interlaced/progressive scan component RGB and single/dual channel monochrome analog video sources; sampling rates up to 30 MHz; three 256x8-bit LUTs connect two RGB or six monochrome video sources; trigger input and timer outputs; 32-bit/33 MHz PCI bus-master; real-time transfer to system or VGA memory.
that such energy may induce, i.e. the Electromagnetic Interference (EMI). The goal of EMC is the correct operation of different electromagnetic equipment avoiding any interference effects. In order to achieve this, EMC pursues two different kinds of issues. Emission issues are related to the unwanted generation of electromagnetic energy by some source, and to the countermeasures which should be taken in order to reduce such generation and to avoid the escape of any remaining energies into the external environment. Immunity issues refer to the correct operation of electrical equipment, referred to as the victim, in the presence of unplanned electromagnetic disturbances. Interference mitigation is achieved by addressing both emission and immunity issues and can be tackled via hardware or software, although a combination of both approaches is usually the best solution. Subsections below describe the minimum measures required to allow the properly working of the camera and an electric motor and driver in the same application, as shown in Fig. 2.

B. Hardware Filtering and Connections

The first action to reduce the interference is to properly connect all the elements used in the application. The equipment’s manufacturer provides important instructions about earth and ground connections that must be followed. In this sense, shield-to-ground connections have to be used. Moreover, EMI filters must be installed with the recommendations provided by the manufacturer, e.g., both filter and driver should be mounted securely on the same conducting earthed backplate. The supply cable should be a stranded conductor to achieve proper connection inside the terminal block. The cable between the driver and the motor should be a screened type and as short as possible with the screen and safety earth core connected directly to the bonded earth post at one end and to the motor earth at the other [10]. The conducting cores should be wound around an output cable filter choke. The paths through which earth current flows have to be clearly define. All earthing leads should be as short as possible and fastened to the bonded backboard earth post. Input and motor output cables should not run together in the same conduit and the control cables to the driver (or other signal cables) should not run alongside motor output cables.

C. Software Filtering

Sometimes not all the EMI can be eliminated by hardware, and software filtering is required to “erase” small particles of image noise [11]. Noise reduction techniques are very similar regardless of the signal being processed, although the characteristics of the expected signal make more suitable one technique or other. Both analogue and digital image signals are susceptible to noise. For example, in Fig. 2 “Video signal” is analogue and it is digitalized by the Matrox card. Noise can be random or white noise introduced by the devices mechanism or processing algorithms.
Operations can be customized with 3 by 3 neighborhoods. However, since opening is the result of MimClose(), opening and closing operations work better in binary images. MIL library provides the functions for opening and closing operations (belonging to the group of morphological operations) that determine each pixel's value according to its geometric relationship with neighborhood pixels. In particular, Low-pass spatial filters are effective in reducing Gaussian frequency and low-frequency components. There are two main types of spatial filters: low-pass filters and rank filters.

Low-pass spatial filters are effective in reducing Gaussian frequency and low-frequency components. There are two main types of spatial filters: low-pass filters and rank filters.

On the other hand, the rank filters are suitable for removing salt-and-pepper type noise since they are not cluttered with shading information. This image processing operation converts an image to a binary image, where the pixel values above the threshold are set to 1, and those below are set to 0. In most cases, it is best to use a rank that is half of the number of elements in the neighborhood, which is called a median filter and that can be performed in MIL library with M_MEDIAN. For example, the median filter is defined as follows:

$$\text{median}(x) = \begin{cases} x_{(i)} & \text{if } 0 \leq i < n/2 \\ x_{(n-i-1)} & \text{if } n/2 \leq i < n \end{cases}$$

For a 3 by 3 neighborhood, the median value is calculated as follows:

$$\text{median}(x) = \begin{cases} x_{(i)} & \text{if } 0 \leq i < 2 \\ x_{(n-i-1)} & \text{if } 2 \leq i < 4 \\ x_{(n-i-1)} & \text{if } 4 \leq i < 6 \\ x_{(n-i-1)} & \text{if } 6 \leq i < 8 \end{cases}$$

where $x_{(i)}$ represents the $i$th element in the neighborhood.

This image processing operation converts an image to a binary image, where the pixel values above the threshold are set to 1, and those below are set to 0. In most cases, it is best to use a rank that is half of the number of elements in the neighborhood, which is called a median filter and that can be performed in MIL library with M_MEDIAN. For example, the median filter is defined as follows:

$$\text{median}(x) = \begin{cases} x_{(i)} & \text{if } 0 \leq i < n/2 \\ x_{(n-i-1)} & \text{if } n/2 \leq i < n \end{cases}$$

For a 3 by 3 neighborhood, the median value is calculated as follows:

$$\text{median}(x) = \begin{cases} x_{(i)} & \text{if } 0 \leq i < 2 \\ x_{(n-i-1)} & \text{if } 2 \leq i < 4 \\ x_{(n-i-1)} & \text{if } 4 \leq i < 6 \\ x_{(n-i-1)} & \text{if } 6 \leq i < 8 \end{cases}$$

where $x_{(i)}$ represents the $i$th element in the neighborhood.

This image processing operation converts an image to a binary image, where the pixel values above the threshold are set to 1, and those below are set to 0. In most cases, it is best to use a rank that is half of the number of elements in the neighborhood, which is called a median filter and that can be performed in MIL library with M_MEDIAN. For example, the median filter is defined as follows:

$$\text{median}(x) = \begin{cases} x_{(i)} & \text{if } 0 \leq i < n/2 \\ x_{(n-i-1)} & \text{if } n/2 \leq i < n \end{cases}$$

For a 3 by 3 neighborhood, the median value is calculated as follows:

$$\text{median}(x) = \begin{cases} x_{(i)} & \text{if } 0 \leq i < 2 \\ x_{(n-i-1)} & \text{if } 2 \leq i < 4 \\ x_{(n-i-1)} & \text{if } 4 \leq i < 6 \\ x_{(n-i-1)} & \text{if } 6 \leq i < 8 \end{cases}$$

where $x_{(i)}$ represents the $i$th element in the neighborhood.

This image processing operation converts an image to a binary image, where the pixel values above the threshold are set to 1, and those below are set to 0. In most cases, it is best to use a rank that is half of the number of elements in the neighborhood, which is called a median filter and that can be performed in MIL library with M_MEDIAN. For example, the median filter is defined as follows:

$$\text{median}(x) = \begin{cases} x_{(i)} & \text{if } 0 \leq i < n/2 \\ x_{(n-i-1)} & \text{if } n/2 \leq i < n \end{cases}$$

For a 3 by 3 neighborhood, the median value is calculated as follows:

$$\text{median}(x) = \begin{cases} x_{(i)} & \text{if } 0 \leq i < 2 \\ x_{(n-i-1)} & \text{if } 2 \leq i < 4 \\ x_{(n-i-1)} & \text{if } 4 \leq i < 6 \\ x_{(n-i-1)} & \text{if } 6 \leq i < 8 \end{cases}$$

where $x_{(i)}$ represents the $i$th element in the neighborhood.
of the two thresholds, the pixel is set to the maximum value in the image; otherwise it is set to 0. The threshold values have to be selected to preserve the required information.

After binarization operation, the original noisy image acquired by the vision system (Fig. 4b) is transformed into a binarized image without noise (see Fig. 4c), after the application of image filtering presented in Fig. 5. This noise is typically due to the electromagnetic interference of the power amplifiers of electric actuators, e.g., AC or DC motors.

C. Erode and Dilate Process

The basic morphological operations [15], erosion and dilation, produce contrasting results when applied to either grayscale or binary images. Erosion shrinks image objects while dilation expands them. Erosion generally decreases the size of the objects and removes small parts by subtracting objects with a radius smaller than the structuring element. With grayscale images, erosion reduces the brightness (and therefore the size) of bright objects on a dark background. With binary images, erosion completely removes objects smaller than the structuring element and removes perimeter pixels from larger image objects.

Dilation generally increases the size of the objects, filling in holes and broken areas, and connecting areas that are separated by spaces smaller than the size of the structuring element. With grayscale images, dilation increases the brightness of objects by taking the neighborhood maximum when passing the structuring element over the image. With binary images, dilation connects areas that are separated by spaces smaller than the structuring element and adds pixels to the perimeter of each image object [16], [17].

Functions MimErode() and MimDilate() of Matrox Image Processing Library (MIL) can be used for basic erosion and dilation operations on 3 by 3 neighborhoods. The noise shown in Fig. 4b is removed after the binarization using these two functions. After this filtering, the Fig. 4c image is obtained.

D. Blob Analysis

The application developed in this work is based on this robust image analysis. Blobs are formed in the image when loops or a configuration of pen strokes get connected into an “ink blot” when the image is closed with a big structuring element [18]. This type of analysis is included in all image processing libraries, e.g., OpenCV, MIL, Sapera, etc. With this package, several important features for the detected blobs can be obtained: area; X and Y location of the centre of gravity; perimeter (normal and convex); size; Feret diameter (minimum, maximum, average) at specific angle; length; breadth; elongation (ratio of length to breadth); shape compactness (closeness of blob particles); number of holes in the blob; central and ordinary moments of any order; etc. For the application developed in this research, the features M_AREA and M_CENTER_OF GRAVITY and the function MBlobSelectFeature() of the MIL library are used. By means of this information, one or several elements (flowers/fruits/vegetables) can be located in the image and in the real world with the camera calibration presented in Section IV.A. This blob analysis is used to reject the sheet shown in Fig. 4c and Fig. 4d images.

E. Computer Program Developed for Image Processing

Fig. 6 presents the flowchart of the developed application. The first step is image acquisition, which includes the software filtering described in Subsection “Erode and Dilate process”. Next, the image is binarized. After that, the opening/closing operations and the erode/dilate algorithms are applied, hence only relevant blobs (i.e., the flowers/fruits/vegetables) remain in the image. The “feature extraction” step obtains the characterization of each detected blobs, i.e. the position of the centre of the flower/fruit/vegetable and the cutting point that allows to separate the element from its peduncle. Finally, the “movement action” step sends the characterization information of each flower/fruit/vegetable to the robotic mechanism responsible for the desired task: recollection, cutting, packaging, classification, fumigation, etc. The core source code of the developed program is shown in Fig. 5, which is included in the paper to facilitate the generation of new specific agricultural applications based on this image processing. This code has been generated in Visual C++ based on MIL library. The image could also be processed using Matlab [19], but in that case the computational cost is higher than using C++ and real-time many not be guaranteed.

The final application or program can be customized to work with different elements (flowers/fruits/vegetables), illumination conditions, element sizes/colours, etc. This flexibility is achieved through 17 configurable parameters, which are available in the “Configuration window”. The most
important are:

1) **IMAGE_THRESHOLD_VALUE**: Threshold (from 0 to 255) used to binarize the acquired image. This value depends on the contrast, brightness, sharpness, etc. of the image.

2) **MIM_BLOB_AREA**: Minimum area to consider an object to be a blob.

3) **MAX_BLOB_AREA**: Maximum area to consider an object to be a blob.

Fig. 8 Examples of applications for flower cutting with a (a) cartesian robot or a (b) pan/tilt robot: the image is processed and the cutting point is obtained.

Fig. 7 Examples of the application performance. The program can be used with many types of flowers/fruits/vegetables, even simultaneously.
to be a blob.

4) **Image BN**: The value is set to 1 if the elements are darker than background (day); otherwise, the value is set to 0 (the elements are clearer than background, i.e. night).

5) **ErodeValue**: Number of pixels used to erode the detected object (flower/fruit/vegetable).

6) **Offset dnm**: Offset (tens of millimetres) added/subtracted to the computed cut position.

V. RESULTS

The goal of the application is to compute the “centre of gravity” (position of the centre) of each detected object (flower, oranges, red pepper, etc.) and the cutting point that allows to separate the element from its peduncle. The computer program processes an acquired image with one element in about 50 milliseconds and processes an image with 8 elements in about 66 milliseconds. This can be considered real time from the processing machine (machine that physically manipulates/processes the elements) point of view.

Fig. 7a shows the appearance of the computer program and the experimental detection of a flower: the centre of gravity is depicted with a white cross and the peduncle cutting point is represented with a black horizontal line. Fig. 7b shows the processed image corresponding to the previous example. Fig. 7c and Fig. 7e show the original image acquired by the camera for other two examples with several oranges and sheets. As before, Fig. 7d and Fig. 7f depict the centre of gravity of each orange with a black cross and the cutting point for collecting them with a white horizontal line. Note that, in both examples the discrimination between oranges and sheets is done properly. The same happens in the examples of Fig. 7g and Fig. 7h (only the processed image is shown) where fruits and vegetables have been combined.

To illustrate the applications of the proposed program, two examples of flower cutting are shown in Fig. 8. In this sense, the program has been customized by the authors for saffron flower cutting giving rise to the commercial software SaffronCut® (2009) (protected by the Spanish office of software), which is currently being used by a Spanish saffron company. In fact, the program described here is intended to be the base for developing new applications that use computer vision in specific agricultural tasks (such as recollection, cutting, packaging, classification, fumigation, etc.) both in the country and in warehouses. That is why this paper emphasizes and solves some important practical problems, such as the EMI, and provides the core source code of the application.

VI. CONCLUSION

In this research, a new vision system to characterize the position of flowers/fruit/vegetables in complex images has been developed. This information can be used for several types of agricultural tasks, such as recollection, cutting, packaging, classification, fumigation, etc. The tests made with the prototype in the laboratory evidence the feasibility and validity of the proposed approach. In particular, the results show that it has good robustness and a very high success. An important characteristic of the proposed algorithm is that it is able to work with several elements inside the camera field of view. This adds flexibility to the application in order to work in the country or in a greenhouse, where the elements are very close to each other and the location of all them must be obtained in real time. We conclude that the use of machine vision in agricultural applications could be increased using the proposed approach, improving the automation of the production and the profits.

REFERENCES


