Jorge Luis García-Alcaraz
Aidé Aracely Maldonado-Macías
Guillermo Cortes-Robles *Editors*

# Lean Manufacturing in the Developing World

## Methodology, Case Studies and Trends from Latin America

Springer

*Editors*
Jorge Luis García-Alcaraz
Departamento de Ingeniería Industrial y Manufactura
Instituto de Ingeniería y Tecnología
Universidad Autónoma de Ciudad Juárez
Chihuahua
Mexico

Guillermo Cortes-Robles
Institute of Technology of Orizaba
Av. Instituto Tecnologico
Orizaba
Mexico

Aidé Aracely Maldonado-Macías
Autonomous University of Ciudad Juarez
Chihuahua
Mexico

ISBN 978-3-319-04950-2 ISBN 978-3-319-04951-9 (eBook)
DOI 10.1007/978-3-319-04951-9
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014934662



Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

# Contents

# Chapter 7
# Automatic Product Quality Inspection Using Computer Vision Systems

**Osslan Osiris Vergara-Villegas, Vianey Guadalupe Cruz-Sánchez, Humberto de Jesús Ochoa-Domínguez, Manuel de Jesús Nandayapa-Alfaro and Ángel Flores-Abad**

**Abstract** The ability to visually detect the quality of a product is one of the most important issues for the manufacturing industry because of the demand made by the consumers is increasing. This process is typically carried out by human experts; unfortunately experts frequently make mistakes because this process could be tedious and tiring even for the most trained operators. A lot of solutions have been proposed to solve this problem, such as the use of lean manufacturing and computer vision systems. This chapter, presents a detailed explanation about the stages involved to create a system to automatically verify the quality of an object using computer vision and digital image processing techniques. First, a revision of state of art researches is presented. This work also focuses on a discussion of the issues involved with computer vision applications. Afterwards, a detailed explanation about the design of two study cases to inspect fabric and apple defects, and its correspondent results are presented. Finally, a point of view about the trends in automatic quality inspection systems using computer vision is offered.

**Keywords** Quality control · Computer vision · Automation

## 7.1 Introduction

Visual inspection is the result of a processing carried out by a part of the brain of the luminous information that arrives to the eyes, and is one of the main data sources of the real world (Sannen and Van Brussel 2012). The information

O. O. Vergara-Villegas (✉) · V. G. Cruz-Sánchez · H. de Jesús Ochoa-Domínguez · M. de Jesús Nandayapa-Alfaro · Á. Flores-Abad
Universidad Autónoma de Ciudad Juárez, Instituto de Ingeniería y Tecnología, Ciudad Juárez, Chihuahua, Mexico
e-mail: overgara@uacj.mx

J. L. García-Alcaraz et al. (eds.), *Lean Manufacturing in the Developing World*,
DOI: 10.1007/978-3-319-04951-9_7, 

perceived with the sense of sight is processed in distinct ways based on the specific characteristics needed for the future tasks to execute. As a result of an image analysis process, the representation of an object is obtained. Immediately a decision is taken to define what to do with the visual information, which typically implies the recognition of the object(s) detected inside a scene and the reactions realized by a body part.

Every day, human beings recognize objects inside a particular scene observed by means of the vision system. This process is done unconsciously (with a minimum effort) even with a lack of the complete knowledge or description of the object to be recognized. Recognition is a term used to describe the ability of human beings to identify the objects around based on previous knowledge (Lee et al. 2010).

The task of visual inspection to recognize objects and to evaluate its quality constitute one of the most important processes in several industries such as the manufacturing industry and alimentary industry, in which, due to the customer demands, it is mandatory to assure the quality of a product (Satorres et al. 2012; Razmjooy et al. 2012; Peng et al. 2008; Sun et al. 2010). The issue of inspecting objects in order to detect defects such color, scratches, cracks or checking surfaces for a proper finish is related to visual quality inspection.

Typically, the quality inspection is made by human experts. However, the experts frequently make mistakes, because the process could be tedious and tiring even for well-trained operators. The problem increases usually due to, the workdays of the inspectors are very long (more than 6 h).

This leads to several industries to look for alternatives to avoid the mistakes made by human inspectors. One of the alternatives adopted by many industries, to remain competitive, is the promoting of lean manufacturing, in which the practices can work synergistically to create a streamlined, high quality system that produces finished products at the pace of customer demand with little or no waste (Sullivan et al. 2002; Abdulmalek and Rajgopal 2007). Unfortunately, the existent evidence suggests that several organizational factors may enable or inhibit the implementation of lean practices among manufacturing plants (Shah and Ward 2003).

Another alternative is to provide to a computer with the ability of inspect and recognize objects automatically. The use of a computer together with other mechanisms, such as cameras, sensors, the knowledge provided by the human expert, and complex algorithms, allows having a capable tool for automatic inspection of the product quality. Then, automation becomes a necessary task for inspection and recognition of objects in order to guarantee the quality of a product.

In this chapter, the issue of the automatic inspection of the quality of an object using computer vision (CV) is addressed. The rest of the chapter is organized as follows.

In Sect. 7.2, a revision of several current designs which use computer vision to inspect the quality of an object is presented. A brief explanation of the issues involved to create a computer vision system using digital image processing techniques is shown in Sect. 7.3. In Sect. 7.4, two study cases are discussed, the first case addressing the problem of fabric color defects detection and the second case addressing the problem of golden apples defect detection. A point of view about the trends in the design of computer vision systems is presented in Sect. 7.5. Finally, Sect. 7.6, presents the conclusions obtained with this work.

## 7.2 Literature Review

In the literature there exist several researches which have addressed the problem of automatic inspection of an object quality using a CV system. Machine vision has been a widely used technology in the industry for the past three decades. It has been an excellent tool for many industrial inspection tasks such as: brake disks, printed circuit board (PCB), float glass, electric contacts, tiles, chickens, fruits, vegetables, fabric, gears, chip alignment, led, to name a few.

The work proposed by (Lerones et al. 2005) defines a solution for the automatic raw foundry brake disk dimensional characterization and visual inspection for the automotive industry. To solve the problem, three CV techniques were used: (a) a calibrated 3D structured-light, for dimensional characterization and inspection, (b) a 3D uncalibrated structured-light, for local fault detection, and (c) a common 2D-vision technique for further local fault recognition. The industrial results show that the described system is appropriate for brake disk dimensional characterization as well as for the detection of hard masses, featheredges, pores, hole jump obstruction, ventilation slot obstruction and veining, providing more efficiency on the production line and bettering working conditions for human operators.

A region-oriented segmentation algorithm for detecting the most common peel defects of citrus fruits is shown in (Blasco et al. 2007). The histogram of an input image is computed to obtain the peak in the lower values corresponding to the background, since it was a uniform black color that contrasted against the orange color of the fruits. Then, the main stage was performed to detect a region of interest consisting of the sound peel, the steam and the defects. This is made by a region growing algorithm followed by a region merging. The algorithm is robust against different varieties and species of citrus fruit and does not need previous manual training or adjustments to adapt the system to work with different batches of fruit or changes in the lighting conditions.

In the chapter presented in (Peng et al. 2008), an online defects inspection method of float glass based on machine vision is presented. The method inspects defects through detecting the change of image gray levels caused by the difference in optic character between glass and defects. Initially, the noise is reduced by means of image filtration based on gradient direction. To remove the back ground stripes, a downward segmentation is implemented. The possible defect core and its distortion are segmented with fixed threshold method and the OTSU algorithm with gray range restrict. Finally, fake defects are eliminated through the method based on defect texture detection. The success of this inspection method provides a reference for defects detection on other materials, such as armor plate.

A CV for the automation of online inspection to differentiate freshly slaughtered wholesome chickens from systemically diseased chickens is presented in

(Yang et al. 2009). The system consisted of a camera used with an imaging spectrograph and controlled by a computer to obtain line-scan images quickly on a chicken processing line of a commercial poultry plant. The system scanned chicken carcasses to locate the region of interest (ROI) of each chicken to extract useful spectra from the ROI as inputs to the differentiation method, and to determine the condition of each carcass as being wholesome or systemically diseased. The high accuracy obtained with the evaluation results showed that the machine vision system can be applied successfully to automatic online inspection for chicken processing. Table 7.1, shows a summary of 11 CV systems contrasting its main features and showing the accuracy recognition rates.

## 7.3 Computer Vision and Digital Image Processing

CV is concerned with modeling and replicating the human vision using computer software and hardware. It is also the main theory for building artificial systems to extract information from images. The main challenge is to combine knowledge in computer science, electrical engineering, mathematics, physiology, biology, and cognitive science in order to understand and simulate the operation of the human vision system. Specifically, CV is a combination of concepts, techniques and ideas from Digital Image Processing (DIP), Pattern Recognition (PR), Artificial Intelligence (AI) and Computer Graphics (CG).

CV describes the automatic deduction of the structure and the properties of a (possible dynamic) three-dimensional world from either a single or multiple two-dimensional images of the world. The computer vision great trick is to extract descriptions of the world from pictures or sequences of pictures (Nalwa 1993).

As a consequence, computer vision systems need the digital image processing techniques to enhance the quality of the acquired images for future use or interpretation. Then, DIP is concerned with taking one array of pixels as input and producing another array of pixels as output, which, in some way, represents an improvement of the original array. DIP is the science of modifying digital images by means of a computer (Forsyth and Ponce 2012).

Frequently, computer vision and digital image processing are erroneously used as a same term. The main difference is in the goals, not in methods. For example, if the goal is to enhance an image for later use it, then this may be called digital image processing, on the other hand, if the goal is to emulate human vision like object recognition, defect detection or automatic driving, then it is closer to computer vision.

There are no clear-cut boundaries in the continuum from DIP at one end to CV at the other. However, one useful paradigm is to consider three types of computerized processes: low-level, mid-level and high-level. Low-level processes involve primitive operations for image preprocessing such as denoising and deblurring, this level is characterized by the fact that both inputs and outputs are images. Mid-level processes involve tasks such as segmentation and description of

**Table 7.1** A summary of 11 computer vision systems to detect defects in objects

| Author (year) | Country | Object | Technique | Defect | Accuracy (%) |
|---|---|---|---|---|---|
| Lerones et al. (2005) | Spain | Brake disk | 3D structured light | Circularity, defective grinding, hard masses, feather edges | 80 |
| Blasco et al. (2007) | Spain | Citrus | Region oriented segmentation | Peel | 95 |
| Peng et al. (2008) | China | Float glass | Texture reverse subtraction | Bubbles, lards, optical distortion | 98 |
| Yang et al. (2009) | USA | Chickens | Thresholding | Wholesome, diseased | 89 |
| Sun et al. (2010) | China | Electric contacts | Blob analysis and PSO | Deckle edge, back cracks, side cracks, eccentricity | 96.7 |
| Mery et al. (2010) | Chile | Corn tortillas | SVM | Size, hedonic scale, production level | 96 |
| Gadelmawla (2011) | Saudi Arabia | Spur gears | Tolerances | Outer diameter, diameter pitch, circular pitch | 99 |
| Perng et al. (2011) | China | LED | Thresholding | Missing components, incorrect orientations | 95 |
| Razmjooy et al. (2012) | Iran | Potato | SVM and ANN | Size, color | 96.86 |
| Možina et al. (2013) | Slovenia | Pharmaceutical tablets | PCA | Debossing shallow | 84 |
| Lin and Fang (2013) | China | Tile | Corner convergence and clustering | Alignment | 90.5 |

individual objects. Finally, the high-level involves making sense of an ensemble of recognized objects and performing the cognitive functions normally associated with vision (Gonzalez and Woods 2009).

As afore mentioned, two broad categories are defined; methods whose inputs and outputs are images and methods whose inputs may be images, but whose outputs are attributes extracted from those images. According to (Gonzalez and Woods 2009) the fundamental stages comprising a CV system for digital image processing are: (a) image acquisition, (b) preprocessing, (c) segmentation, (d) feature extraction (representation and description), (e) Recognition (interpretation and classification), and (f) knowledge base.

In Fig. 7.1 a graphical representation of each stage and the mode of interaction of each with the other are shown. Additionally, in the following subsections a brief explanation of each stage is presented.

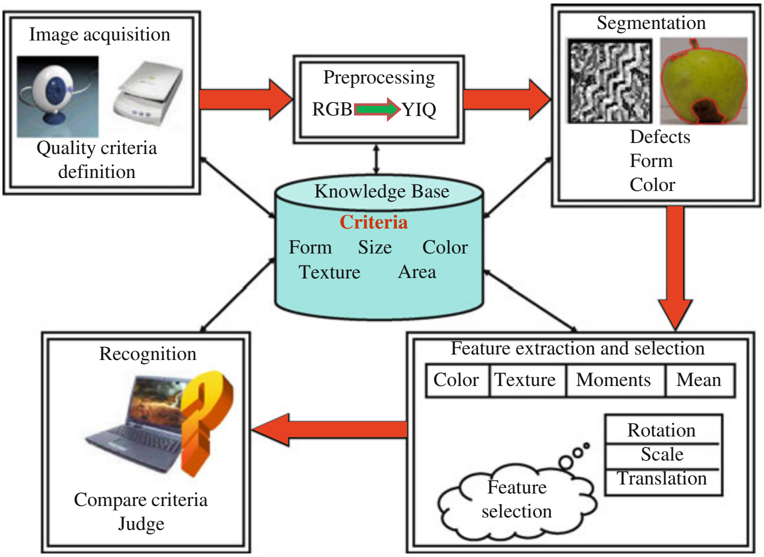


**Fig. 7.1** Fundamental steps of a computer vision for digital image processing

## 7.3.1 Image Acquisition

Before any image processing can start, an image must be captured by a camera and converted to a manageable entity. Thus, in order to acquire a digital image, an image sensor and the ability to digitize the signal produced by that sensor are needed (Wandell et al. 2002). The sensor can be a television camera, a line scan camera, video, scanner, etc. If the output of the sensor is not digital, then an analog to digital converter is necessary to digitize the image. The digital image is obtained as a result of sampling and quantization of an analog image or created already in digital form.

Typically, a digital image is represented as a bi-dimensional matrix of real numbers. The convention $f(x,y)$ is used to refer an image with size $M \times N$, where $x$ denote the row number, and $y$ the column number. The value of the bi-dimensional function $f(x, y)$ at any pixel of coordinates $(x0, y0)$, denoted by $f(x0, y0)$, is called the intensity or gray level of the image at that pixel.

## 7.3.2 Preprocessing

After a digital image has been acquired, several preprocessing methods can be applied in order to enhance the data of the image prior to the computational processing. In this stage the image is processed and converted into a suitable form for further analysis (Choi et al. 2011). Most of the computer vision applications require taking care in designing the processing stage in order to achieve acceptable results. Preprocessing operations are also called filtration.

Examples of such operations include smoothing, exposure correction and color balancing, noise reduction (denoising), increasing sharpness, image deblurring, image plane separation, normalization, etc. The image obtained after this stage is the input to the segmentation step.

### *7.3.3 Segmentation*

Autonomous segmentation is one of the most difficult stages in DIP. Segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as superpixels). The goal is to simplify and to change the representation of an image into something that is more meaningful and easier to analyze (Padma and Sukanesh 2013). A coarse segmentation process delays the outcome of a satisfactory solution in a DIP problem. On the other hand, a weak segmentation process in most of the cases will lead to errors.

As a result of the segmentation process, the raw pixel data that constitute the boundaries among the regions or information about which pixel belongs to which region is obtained. The image segmentation is the operation that marks the transition between low-level to mid-level image processing. Among the most commonly used segmentation methods are: thresholding, contour detection, region based, morphological watersheds and region growing.

### *7.3.4 Feature Extraction*

Once an image has been segmented, the resulting individual regions can be described. Feature extraction, also called image representation and description, is the operation performed to extract and highlight features with some quantitative information which is essential to distinguish one class of objects from another. It is a critical step in most computer vision solutions because it marks the transition from pictorial to non pictorial data representation.

In order to store the characteristics extracted from an object an $n \times 1$ array called feature vector is built. The feature vector is a compact representation of an image and its content can be symbolic, numerical or both. The main challenge in this step is that the features extracted must be invariant to changes in rotation, scale, translation and contrast. Obtaining the invariants ensures that the computer vision system will be able to recognize objects even when they appear with different contrast, size, position and angle inside the image (Mullen et al. 2013).

The important features extracted include points, straight lines, regions with similar properties, color, textures, shapes, and a combination. The boundary descriptors such as statistical moments or Fourier descriptors are often used to tackle the problem of invariance.

### *7.3.5 Recognition*

This stage constitutes the high-level of image processing. Recognition is the process to assign a label to an object, based on the information provided by its descriptors. Moreover, it involves assigning meaning to a set of recognized objects (Xiao and Wang 2013). The recognition algorithms analyze the numerical properties of various image features and classify data into categories. This stage is composed of two phases: training and testing.

All the classification algorithms are based on the assumption that the image in question depicts one or more features and that each of these features belongs to one of several distinct and exclusive classes. The classes may be stated a priori by an analyst specifying the number of desired categories (supervised classification) or automatically clustered (unsupervised classification) into sets of prototypes.

Quite often, the algorithms used for recognition include artificial neural networks (ANN), support vector machines (SVM), distance and similarity measures (k-Nearest Neighbors, Bayesian, Euclidean, Manhattan), and matching.

### *7.3.6 Knowledge Base*

The a priori knowledge about a specific image processing problem is coded in the form of a knowledge database. The database may be as simple as detailing regions of an image where the information of interest is known to be located, thus limiting the search that has to be conducted for seeking information; or can be quite complex such as inter related list of all major possible defects in a material inspection problem (Gonzalez and Woods 2009). The knowledge base guides the operation of each processing module and also controls the interaction among all the modules.

## 7.4 Study Cases

As described in Sect. 7.2, several computer vision systems have been created in different industries. Even, when almost all the systems are build based on the steps explained in Sect. 7.3, not always all the steps are used. Sometimes, other more specific steps are included. The purpose of the current section is to present two study cases, one is to offer the specific details on how to build a computer vision system using digital image processing techniques to inspect the color quality fabric, and the other to inspect defects in apples.

## *7.4.1 Automatic Inspection of Fabric Quality*

In the textile industry, defects detection is of crucial importance in the quality control of manufacturing (Bissi et al. 2013). The common inspection process is made by human experts, but since this task is dull and repetitive many errors are frequently committed. It has been estimated that the price of fabrics is reduced from 45 to 65 % due to the presence of defects (Kumar 2008).

The fabric is affected by yarn quality, loom defects and stamping process. The textile industry has the necessity to verify the quality of a stamping fabric process. The following subsections explain the design and implementation of a computer vision system to inspect color degradation in fabric.

### 7.4.1.1 Definition of the Quality Criteria

In the textile industry a common error that causes defects in the stamping process of a fabric roll is that some of the colors being used goes finishing. For a human inspector, this color degradation can be detected until the color difference between the original and the degraded fabric is very noticeable. For this example, the quality of a fabric is defined by the conservation of colors at the stamping process.

The input to the system is a criterion about the quantity of color variations accepted to consider a fabric without defects (good quality). At this stage a range between 1 and 10 % of color degradation is accepted for a good quality fabric, while out of that range the fabric is considered of bad quality.

### 7.4.1.2 Acquisition of Fabric Images

The stamping process was simulated using artificial fabric images, due to real stamping process imaging could not be achieved. Three databases of images were selected, each one containing what is called a texton that is the main element that compose a texture. The public databases used were: “Artificial Color Texture”, “Forrest textures”, and “Texture Gallery”. A subset of ten textons was selected from each database.

### 7.4.1.3 Preprocessing of Fabric Images

To automatically simulate the creation of a fabric roll, each texton was repeated one after another 60 times. This operation was made randomly with rotations of 90

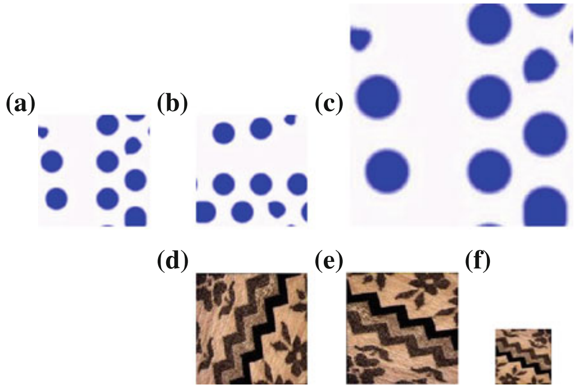


**Fig. 7.2** Examples of textons. **a** Original texton: 125_exhaustive_p1, **b** 125_exhaustive_p1 rotated 90° clockwise, **c** 125_exhaustive_p1 doubled, **d** Original texton: S_S_cloth2, **e** S_S_cloth2 rotated counterclockwise, and **f** S_S_cloth2 rotated clockwise and halved

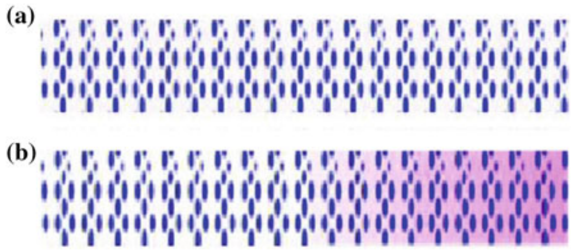


**Fig. 7.3** Examples of fabric rolls. **a** Good quality fabric, and **b** Bad quality fabric

and 180° clockwise, scaling of doubled and halved of the original size, scaled and rotated together, and with addition of 10 and 40 % of salt and pepper noise with zero mean and a variance of 400. A random degradation in red, green and blue (RGB) colors of 1, 2, 5, 10, 20, 40 and 50 % was applied.

At the end of this process 1400 images were created, 700 images of good quality fabric and 700 images of bad quality fabric. In Fig. 7.2, several textons with different changes are depicted. In Fig. 7.3 an example of good quality fabric and bad quality fabric is shown.

#### 7.4.1.4 Feature Extraction from Fabric Image Rolls

At this stage, a feature vector to store the statistical features associated with each fabric roll was computed. The process of the texture feature extraction was carried out in three RGB (Red, Green, Blue) planes and in the intensity plane of the HSI (Hue, Saturation, Intensity) model. In summary, the features extracted to handle image invariance were: 10 normal moments, 10 central moments, 7 Hu moments, 6 Sidharta Maitra moments, mean, variance and standard deviation in red, green and blue, and mean for the intensity of the HSI model. At the end of this step, a feature vector with 43 characteristics was obtained for each fabric roll.

#### 7.4.1.5 Feature Selection

This stage consists on the selection of the more discriminate features to distinguish one fabric from another. In order to perform this stage, the so called testor theory was used. The algorithm used was the well known BT algorithm to obtain typical testors.

The BT algorithm defines a testor for two classes $T_0$ and $T_1$ as: "A set $t = (i1,\ldots, is)$ of columns of the table $T$ (and the respective features $xi1,\ldots, xis$), is called testor for $(T_0, T_1) = T$, if after eliminating from $T$ all of the columns except those belonging to $t$, does not exist any row in $T_0$ equal to $T_1$".

A testor is called irreducible (typical) if upon eliminating any columns it stops being testor for $(T_0, T_1)$, where $s$ is called the length of the testor. The BT algorithm for typical testors is based in the idea of generate boolean vectors of $\alpha = (0,\ldots.., 0, 1)$ until arrive to $\alpha = (1,1,\ldots.,1, 1)$. For each case, it is verified if the set of columns that correspond to the coordinates of the $n$-tuple generated is a typical testor.

At this point, it is needed to have the learning matrix which contains the descriptions of objects in terms of a set of features; the matrix of differences, which it is obtained from the learning matrix comparing the values of object features pertaining to different classes; and the basic matrix, that is formed exclusively by basic rows (incomparable rows).

After computing the BT algorithm, a measurement of the importance of a single feature was calculated. After the typical testors were obtained, an irreducible combination of fewer features was obtained, each feature was very important to assure the difference from one class to another (bad vs good quality). Therefore, if a feature appears in many testors, it cannot be eliminated because it helps to preserve the class separation.

After this phase, the feature vector was reduced by a 76.7 %. Only 10 features were selected from the original set of 43 features to be used in the final recognition step. The features selected were: 2 central moments, 3 Maitra moments, red and blue color mean, color green variance, illumination mean and blue standard deviation.

#### 7.4.1.6 Recognition of Fabric Images

In the recognition step, the voting algorithm (Alvot), which is based on partial precedence or partial analogies, was used. The premise of the algorithm is the following: if an object may looks like another, but not totally, and that the parts that look alike can offer information about possible regularities. Thus, a final decision is taken. The model of voting algorithms is described in six steps: (a) obtain the system of support sets, (b) computation of the similarity function, (c) evaluation by row for a given support set, (d) evaluation by class for a given support set, (e) evaluation by class for all the system of support sets, and (f) obtaining the solution rule.

After the computation of all the stages involved in voting algorithm, the discrepancies among the object inspected and its correspondent model were determined and a decision about the membership to good or bad quality class of the inspected object is emitted.

#### 7.4.1.7 Experimentation and Results

Five different types of tests were carried out: *Case* (*a*) *Validation of the system learning ability:* Consists on validating if the system performs the training in a correct way. In these test, the system does not fail when an image of the training set is used as input. *Case* (*b*) *Rotation changes:* The goal is to verify if the system can handle rotation effects at 90 and 180° with regard to the original fabric images. *Case* (*c*) *Scale changes:* The goal is to verify if the system can handle scaling effects of halved and doubled the size of the original fabric images. *Case* (*d*) *Rotation and scale changes:* This test was made to verify the case b and c together. *Case* (*e*) *Noise insertion:* The objective is to evaluate if the system can handle images contaminated with salt and pepper noise, with a probability of 10 % for good quality and 40 % for bad quality.

The results obtained in each case are depicted in Table 7.2.

As can be seen in Table 7.2, the techniques utilized were effective and an average recognition of 71.2 % was obtained. The main contribution is that the system can handle invariance to rotation, scale and noise in a separated and in a combined way.

### *7.4.2 Automatic Inspection of Apple Quality*

The creation of computer vision applications for agricultural industry has increased considerably in recent years (Cubero et al. 2011). The agricultural industry has the necessity to verify the quality of fruits and vegetables during the storage process. Typical target applications of inspection systems include grading and quality estimation from external parameters or internal features. Much of the sorting and grading processes are still not automatic. Manual inspection of fruit is tedious and can cause eye fatigue. Frequently, the inspection process is subject to errors due to different judgment emitted by different persons.

This study case, addresses the problem of design and implement a computer vision system for the automatic inspection of the quality fruit, particularly the golden apples. Apples are very susceptible to damage and the presence of bruises on the apple skin affects not only the appearance of the apple, which is an important indicator of quality, but also accelerates its deterioration (Santos and Rodrigues 2012). Two main problems were detected for the implementation of a computer vision system for apple grading: (1) how to acquire the whole surface

**Table 7.2** Results obtained for fabric roll color inspection

| Case | # images | % of success | % of failure |
|---|---|---|---|
| a | 200 | 100 | 0 |
| b | 360 | 69.7 | 30.3 |
| c | 360 | 64.15 | 35.85 |
| d | 240 | 72.31 | 27.69 |
| e | 240 | 50.2 | 49.8 |
| Total | 1400 | 556.36 | 143.64 |
| Average | | 71.2 | 28.8 |

image of an apple by cameras at an on-line speed, and (2) how to quickly identify the stem, calyx and the presence of defects (Li et al. 2002).

The solution presented for apple classification tackles both problems. Additionally, the system not use only the typical numerical information obtained from a computer vision system but also a symbolic knowledge obtained from a human expert is added to enhance the ability of the system to classify the apples. The approach presented is called a neuro-symbolic hybrid system (NSHS).

#### 7.4.2.1 Definition of the Quality Criteria

A hybrid system offers the possibility to integrate two or more knowledge representations of a particular domain in a single system. One of the main goals is to obtain the complementary knowledge that allows improving the efficiency of the global system. A specific example of a hybrid system is the so called NSHS, which is mainly based on the symbolic representation of an object obtained from a human expert in the form of rules and a computer vision system to obtain the numerical information.

The quality criteria to evaluate a golden apple were obtained from a human expert in apple classification. For the case of golden apples, a category is assigned depending on the value of the external attributes. There exist four categories: category extra, category I, category II, and category III. In this example, only the category extra is evaluated, by that an apple can belong to one of two classes: good or bad quality. In Fig. 7.4, an example of good and bad quality apples are depicted. Additionally, in Table 7.3 a resume of the external attributes of a golden apple with its associated variable name, as well as the value and type are shown.

#### 7.4.2.2 Image Acquisition

For the problem of category extra golden apple inspection, 148 images were obtained by means of a digital camera. The golden apples were acquired from crop fields at Puebla city in Mexico. For the complete set of images an operation of rotating 90 and 180° clockwise, doubled and halved the scaling and noise addition, similar to the study case of fabric was performed. At the end of this process the set

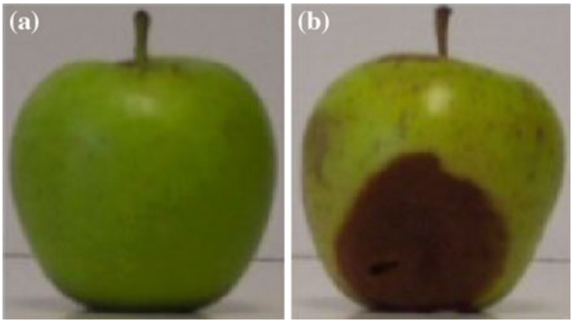


**Fig. 7.4** Examples of category extra golden apples. **a** Good quality apple, and **b** Bad quality apples

**Table 7.3** Quality criteria determined to measure the quality of an apple

| Attribute | Acronym | Type | Value |
|---|---|---|---|
| Lengthened defect | LD | Range | 0–6 |
| Spotted defect | SD | Range | 0–2.5 |
| Various defects | VD | Range | 0–5 |
| Stem | S | Binary | True/false |
| Red color | RC | Range | 0–255 |
| Green color | GC | Range | 0–255 |
| Blue color | BC | Range | 0–255 |

of 148 apples were divided into two categories: bad (74) and good (74) quality. In Fig. 7.5, an example of different apples after changing in rotation and scale are depicted.

#### 7.4.2.3 Image Preprocessing

The stage consists of the image conversion from RGB color model to YIQ (Luminance, Inphase, Quadrature) color model. The main reason to perform such conversion is to facilitate the image feature extraction. The YIQ model was computed to separate the color from the luminance component because of the ability of the human visual system to tolerate more changes in reflectance than to changes in shade or saturation. The main advantage of the model is that reflectance (Y) and the color information (I and Q) can be processed separately. Reflectance is proportional to the amount of light perceived by human eye.

#### 7.4.2.4 Apple Feature Extraction

The characteristics of each image were extracted based on the information defined by the human expert in the form of rules and by means of image processing in the

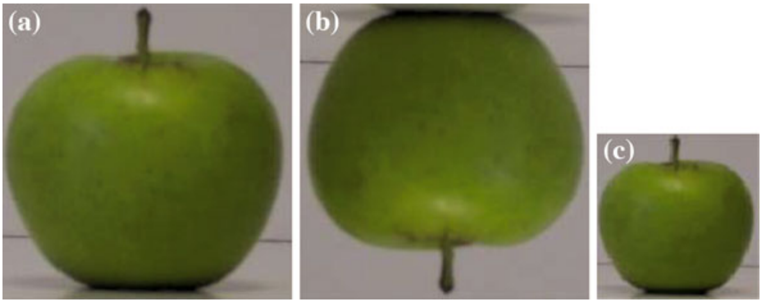


**Fig. 7.5** Golden apples of category extra. **a** Original apple image, **b** Apple rotated 180° clockwise, and **c** Apple halved the original size

form of numerical data. These two types of knowledge information were combined in order to obtain the overall representation of an apple.

The number of rules defined by the experts were four, an example of a single rule is the following: *"IF an apple has the corresponding color, and has the stem, and has lengthened defects that do not exceed 2 cm, and has several defects that do not exceed 1 cm$^2$, and has spotted defects that do not exceed ¼ cm$^2$, THEN the apple belongs to category extra with good quality".*

In order to obtain the numerical values, the same process of feature extraction for the problem of fabric was performed.

At the end of this stage, the rules were compiled with a knowledge based artificial neural network (KBANN), in order to obtain a representation of the information that further can be combined with the numerical results extracted from the computer vision system. The combination was made using the method called Neusim, which is based on the cascade correlation Fahlman algorithm.

#### 7.4.2.5 Apple Classification

The output of the apple feature extraction stage was a joined representation of the symbolic and numerical knowledge. In order to further classify a golden apple, a refinement of that knowledge must be made. That refinement is performed by running again the Neusim method, but now not to join knowledge but use it as a classifier.

The main advantage of use Neusim algorithm is that one can see the number of hidden units added during the learning process, this is very useful to monitor the complete process of incremental learning. The output of this stage is the decision about the quality of an apple in one of two classes, bad or good.

#### 7.4.2.6 Experiments and Results

From the total set of 128 golden apple images, 74 were used for the stage of training and 74 for the stage of recognition. In order to carry out the experiments, three different approaches were selected: (a) a connectionist approach which uses only the data obtained from the computer vision system, (b) a symbolic approach which uses only the data obtained from the compiled rules, and (c) a NSHS, which is a combination of connectionist and symbolic approach.

For the case of the tests using the connectionist approach, three scenarios were defined: (a) using the numerical data obtained from the overall 148 images (100 %), (b) using the data obtained only from 111 images (75 %), and (c) using only data from 74 images (50 %). Three rules were used to obtain the results referring to the test case using the symbolic approach. The first rule called 7 involves the following seven attributes: LD, SD, VD, S, RC, GC, and BC. The second rule called R5 considers the following five attributes: RC, GC, BC, S, and LD. Finally, the third rule called R4 includes the following four attributes: LD, SD, VD, and S. For the case of NSHS approach, a combination of connectionist and symbolic was made. Three rules R7, R5, and R4 were combined with 100, 75 and 50 % of the total number of examples. The overall results obtained are shown in Table 7.4.

One of the typical problem causing failures in computer vision systems is the lack of the complete description of an object. This can be observed by verifying the results obtained in Table 7.4 with symbolic and connectionist approaches. This disadvantage can be withdrawn by using a method to complete the information with the data defined by the knowledge of a human expert. The systems which allow this combination types are called NSHS, as can be seen with results shown in Table 7.4; these systems are very efficient for complementing the necessary knowledge for an automatic object inspection by means of a computer vision system.

For example, in the pure symbolic approach, the rule R4 was not enough to classify the apples correctly, but when it is integrated with the group of numeric examples (100, 75, 50 %), a substantial improvement is obtained, because the knowledge that does not contain the rule, is complemented with the numerical base of examples.

## 7.5 The Future of Computer Vision Systems

Even when the last few years have shown many computer vision systems which have been proven to be very efficient for the particular task for which they were created, there are still many challenges to be solved. The researchers divide these challenges into two categories, issues referring to hardware design techniques and issues concerned to software algorithms (Andreopoulos and Tsotsos 2013).

**Table 7.4** The results obtained for the three approaches proposed for tests

| Approach | Compiled rules | % of examples used | Accuracy (%) |
|---|---|---|---|
| Connectionist | – | 100 | 95.14 |
| | – | 75 | 91.21 |
| | – | 50 | 90.54 |
| Symbolic | R7 | – | 93 |
| | R5 | – | 90.12 |
| | R4 | – | 14.19 |
| NSHS | R7 | 100 | 96.62 |
| | R7 | 75 | 95.27 |
| | R7 | 50 | 90.54 |
| | R5 | 100 | 95.27 |
| | R5 | 75 | 95.94 |
| | R5 | 50 | 96.62 |
| | R4 | 100 | 91.22 |
| | R4 | 75 | 93.24 |
| | R4 | 50 | 94.59 |

### *7.5.1 Issues Concerning Hardware*

This challenge concerns the design and build of specific hardware to solve the typical problems of computer vision, due to that there is a significant gap in terms of the input size and the computational resources needed to reliably process those inputs. The solution needs to take into account variables such as real time acquisition huge storage space, distributed information, depth information, parallel execution, and portability. Following several of the hardware trends are explained.

*Designing of sensors such Microsoft Kinect*: With the invention of the low-cost Microsoft Kinect sensor, high-resolution depth and visual RGB sensing has become available for widespread uses. The complementary nature of the depth and visual information provided by the Kinect sensor opens up new opportunities to solve fundamental problems in computer vision and to design and build other powerful sensors based on this technology. The main topics that will be solved include preprocessing, object tracking and recognition, human activity analysis, hand gesture analysis, and indoors 3-D mapping.

*Distributed smart cameras*: This implies the design and implementation of real time distributed embedded systems that perform computer vision using multiple cameras. This approach has emerged thanks to a confluence of simultaneous advances in disciplines such: computer vision, image sensors, embedded computing, and sensor networks. Processing images in a network of distributed smart cameras introduces several complications such as the tracking process. The distributed smart cameras will represent key components for future embedded computer vision systems and that smart cameras will become an enabling technology for many new applications.

*Multicore processors*: The recent emergence of multi-core processors enables a new trend in the usage of computers. Computer vision applications, typically requires heavy computation and lots of bandwidth, which makes difficult to run the applications in real time. The use of multi-core processors can potentially serve the needs of such workloads. In addition, more advanced algorithms can be developed utilizing the new computation paradigm. The main advantage offered will be the execution of parallel algorithms, for example by taking multiple cameras inputs of a scene, extracting useful features, and performing statistical inference almost at the same time. In resume, the fact of parallelizing the workload will speed up the performance of the computer vision systems.

*Graphics processing units* (*GPU*): Computer vision tasks are computationally intensive and repetitive, and they often exceed the capabilities of the CPU, leaving little time for higher level tasks. However, many computer vision operations map efficiently onto the modern GPU, whose programmability allow a wide variety of computer vision algorithms to be implemented. The GPU provides a streaming, data parallel arithmetic architecture which carries out a similar set of calculations on an array of image data. The single instruction multiple data (SIMD) capability of the GPU makes it suitable for running tasks, which often involve similar calculations operating on an entire image.

*Mobile devices*: Since mobile devices are gaining increasingly powerful processors, great quality cameras and high speed network access, they will be suitable to implement computer vision algorithms. The ubiquitous platform offered by a mobile device can be used for a wide range of applications such as: smart surveillance, assembly line troubleshooting aid, controlling specific devices with gestures or screen touches, and entertainment. Additionally, the sensors often included in mobile devices such as: GPS, gyroscopes and accelerometers can contribute to enhance the computer vision tasks. The main disadvantage is the life time battery.

### 7.5.2 Issues Concerning Software

One of the main problems to solve is to find a way to generalize the knowledge as well as the human vision system, enhance the way for the representation of objects, and the ability for learning and inferencing the models used. Following several of the software trends are explained.

*Multispectral and hyper spectral image analysis*: The data of the images is obtained at specific frequencies across the electromagnetic spectrum. The spectral imaging allows the extraction of additional information which human eye fails to capture with its receptors for red, green and blue. After the representation in the electromagnetic spectrum, the images are combined to form a three-dimensional hyper spectral data cube for processing and analysis. The main difference between multispectral and hyper spectral images is the number of bands obtained. The main areas for the application of this type of imaging are agriculture, mineralogy, physics, and surveillance.

*RGB-D images for representation and recognition*: This type of images are formed with the classical red, green and blue information and adding information about the scene depth, which is obtained by a technique called structured light implemented with an infra red laser emitter. This type of images is obtained using a kinect sensor and the data containing visual and geometric information of the scene. This image offers the advantage to obtain the depth information using only one device instead of using the classical pair of images. The features obtained with RGB-D images will be very useful to improve the process of shape estimation, 3D mapping and localization, path planning, navigation, pose recognition and people tracking.

*Real time display of obscured objects*: This kind of algorithms could help car drivers and airplane pilots to see through fog, and submarines explore under the sea. The algorithms could provide safety features for future intelligent transportation systems. This means that computer vision software modules can distinguish specific objects more accurately from the rest of a scene, even in complete darkness.

*Deep and feature learning*: One of the main challenges for computer vision and machine learning is the problem of learning representations of the perceptual world. The learning and perceptual methods allows automatically learn a good representation of the input unlabeled data, offering a time reduction against typical learning algorithms which spend a lot of time to obtain the input feature representation. Since these algorithms mostly learn from unlabeled data, they have the potential to learn from vastly increased amounts of data (since unlabeled data is cheap), and therefore also achieving a vastly improved performance. The representation of multilevel hierarchies obtained are useful not only for low level features such as edge or blob detectors, but also are useful for high level concepts such as face recognition.

*Mid-level patch discovery*: The technique allows discovering a set of discriminative patches that can serve as a fully unsupervised mid-level visual representation. The desired patches need to satisfy two requirements (1) to be representative, they need to occur frequently enough in the visual world, and (2) to be discriminative, that is, they need to be different enough from the rest of the visual world. The patches could correspond to parts, objects, visual phrases, etc. but they are not restricted to be any one of them. The patches are simple to compute, and offers very good discriminability, broad coverage, better purity, and improved performance compared to visual world features. These approaches can be used mainly on scene classification, beating bag-of-words, spatial pyramids, object bank, and scene deformable-parts models.

*Augmented reality*: Allows enhancing the senses of a user by manipulating virtual objects superimposed on top of the real world scenes. In other words, AR bridges the gap between the real and the virtual in a seamless way. Specifically, because of the improved ease-of-use of augmented reality interfaces, these systems may serve as new platforms to gather data, as imagers may be pointed by users to survey and annotate objects of interest to be stored in different kind of systems such as educational learning, games, training skills and geographical information systems (GIS).

*Gesture control*: Human–computer interfaces (HCI) needs to evolve from mouse-keyboard based interaction, multi-touch screens and other exotic approaches such as using special gloves or other devices, until reach the use of gestures to translate human actions into control applications. There were and still are a series of attempts to produce computer control scripts via gesture-based interfaces, and the research literature is abundant in chapters on this subject. The main challenge is that the computer-based control via gestures has to be robust and in real-time. Any lag in the result can lead to users abandoning it.

## 7.6 Conclusions

In this chapter the design and implementation of two computer vision systems to verify the quality of fabric and apples was presented. The main challenges to build the computer vision systems were explained exhaustively. A brief explanation and contrasting of several computer vision systems already presented in the literature were discussed. Additionally, a highlighting was made about its ability or accuracy to detect defects of an object. Finally, a brief explanation about the future of computer vision systems for quality inspection was presented.

Despite the numerous studies developed in the computer vision area, there is still not a standardized method which could be proposed for the assessment of the quality of different types of objects. The particular characteristics of an object require a computer vision system to be customized; this implies an exhaustive research process and not only the purchase of expensive equipment. Besides, in order to obtain a better performance of the system, the acquisition of the knowledge from the human experts and the techniques to represent it in terms of numerical information is mandatory.

## References

Abdulmalek, F., & Rajgopal, J. (2007). Analyzing the benefits of lean manufacturing and value stream mapping via simulation: A process sector case study. *International Journal of Production Economics, 107*(1), 223–236.

Andreopoulos, A., & Tsotsos, J. (2013). 50 years of object recognition: Directions forward. *Computer Vision and Image Understanding, 117*(8), 827–891.

Bissi, L., Baruffa, G., Placidi, P., Ricci, E., Scorzoni, A., & Valigi, P. (2013). Automated defect detection in uniform and structured fabrics using Gabor filters and PCA. *Journal of Visual Communication and Image Representation, 24*(7), 838–845.

Blasco, J., Aleixos, N., & Moltó, E. (2007). Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm. *Journal of Food Engineering, 81*(3), 535–543.

Choi, J., Ro, Y., & Plataniotis, J. (2011). A comparative study of preprocessing mismatch effects in color image based face recognition. *Pattern Recognition, 44*(2), 412–430.

Cubero, S., Aleixos, N., Moltó, E., Gómez-Sanchis, J., & Blasco, J. (2011). Advances in machine vision applications for automatic inspection and quality evaluation of fruits and vegetables. *Food and Bioprocess Technology, 4*(4), 487–504.

Forsyth, D., & Ponce, J. (2012). *Computer vision: A modern approach* (2nd ed.). New Jersey, USA: Pearson.

Gadelmawla, E. (2011). Computer vision algorithms for measurement and inspection of spur gears. *Measurement, 44*(9), 1669–1678.

Gonzalez, R., & Woods, R. (2009). *Digital image processing* (3rd ed.). New Jersey, USA: Prentice Hall.

Kumar, A. (2008). Computer-vision-based fabric defect detection: A survey. *IEEE Transactions on Industrial Electronics, 55*(1), 348–363.

Lee, S., Kim, K., Kim, J., Kim, M., & Yoo, H. (2010). Familiarity based unified visual attention model for fast and robust object recognition. *Pattern Recognition, 43*(3), 116–1128.

Lerones, P., Fernández, J., García-Bermejo, J., & Zalama, E. (2005). Total quality control for automotive raw foundry brake disks. *International Journal of Advanced Manufacturing and Technology, 27*(3–4), 359–371.

Li, Q., Wang, M., & Gu, W. (2002). Computer vision based system for apple surface defect detection. *Computers and Electronics in Agriculture, 36*(2–3), 215–236.

Lin, K., & Fang, J. (2013). Applications of computer vision on tile alignment inspection. *Automation in construction, 35*(November), 562–567.

Mery, D., Chanona-Pérez, J., Soto, A., Aguilera, J., Cipriano, A., & Veléz-Rivera, N., et al., (2010). Quality classification of corn tortillas using computer vision, *Journal of food engineering*, 101(4): 357–364.

Možina, M., Tomaževic, D., Pernuš, F., & Likar, B. (2013). Automated visual inspection of imprint quality of pharmaceuticals tablets. *Machine Vision and Applications, 24*(1), 63–73.

Mullen, R., Monekosso, D., & Remagnino, P. (2013). Ant algorithms for image feature extraction. *Experts Systems with Applications, 40*(11), 4315–4332.

Nalwa, V. (1993). *A guided tour of computer vision* (1st ed.). Boston, Massachusetts, USA: Addison-Wesley Longman Publishing Co.

Padma, A., & Sukanesh, R. (2013). Wavelet statistical texture features-based segmentation and classification of brain computed tomography images. *IET Image Processing, 7*(1), 25–32.

Peng, X., Chen, Y., Yu, W., Zhou, Z., & Sun, G. (2008). An online defects inspection method for float glass fabrication based on machine vision. *International Journal of Advanced Manufacturing Technology, 39*(11–12), 1180–1189.

Perng, D., Liu, H., & Chang, C. (2011). Automated SMD LED inspection using machine vision. *The International Journal of Advanced Manufacturing Technology, 57*(9–12), 1065–1077.

Razmjooy, N., Somayeh, B., & Soleymani, F. (2012). A real-time mathematical computer method for potato inspection using machine vision. *Computers and Mathematics with Applications, 63*(1), 268–279.

Sannen, S., & Van Brussel, H. (2012). A multilevel information fusion approach for visual quality inspection. *Information Fusion, 13*(1), 48–59.

Santos, J., & Rodrigues, F. (2012). Applications of computer vision techniques in the agriculture and food industry: A review. *European Food Research and Technology, 5*(6), 989–1000.

Satorres, S., Gómez, J., Gámez, J., & Sánchez, A. (2012). A machine vision system for defect characterization on transparent parts with non-plane surfaces. *Machine Vision and Applications, 23*(1), 1–13.

Shah, R., & Ward, P. (2003). Lean manufacturing: context, practice bundles, and performance. *Journal of Operations Management, 21*(2), 129–149.

Sullivan, W., McDonald, T., & Van Aken, E. (2002). Equipment replacement decisions and lean manufacturing. *Robotics and Computer Integrated Manufacturing, 18*(3–4), 255–265.

Sun, T., Tseng, C., & Chen, M. (2010). Electric contacts inspection using machine vision. *Image and Vision Computing, 28*(6), 890–901.

Wandell, B., Gamal, A., & Girod, B. (2002). Common principles of image acquisition systems and biological vision. *Proceedings of the IEEE, 90*(1), 5–17.
Xiao, B., & Wang, G. (2013). Generic radial orthogonal moment invariants for invariant image recognition. *Journal of Visual Communication and Image Representation, 24*(7), 1002–1008.
Yang, C., Chao, K., & Kim, M. (2009). Machine vision system for online inspection of freshly slaughtered chickens. *Sensors and Instruments for Food Quality and Safety, 3*(1), 70–80.