Machine Vision Application to Automatic Detection of Living Cells/Objects

Hernando Fernández-Canque
Glasgow Caledonian University, Scotland,
UK

1. Introduction

The concept of Machine Vision (MV) originates in the early 1950's and practical MV applications appeared in the early 1980's. Early theoretical work suggested serious limitations in computational abilities as the main reason for inefficient use of MV. MV is a 'simple' processing system which receives and combines signals from cameras by manipulating images at the pixel level to extract information that can be used in a decision making activity to produce the output required. MV has several advantages over systems utilising conventional technologies. The attention span of human operators is relatively short (Butler, 1980). Camera images are detailed and therefore contain a large amount of information. This combined with the computer power available provides a huge potential for MV applications. The chapter will provide an overview on the developments and historical evolution of the concept of MV applications. This chapter will concentrate on MV application to automatic detection of objects with ill defined shape, size, and colour with high variability. Objects that change with time have no fixed structure; can present extra problems on application of automatic detection using MV. For this type of application, current manual detection requires highly specialised and highly trained operators. The application of MV will facilitate the detection of these objects with the advantage of fast response. It is easy to use, cost effective with consistent and reliable results. The detection of micro-organisms and the detection of suspicious activity in humans fall into this category. The first example examines development of an automatic system for microscopic examination of the recovered deposit for the detection and enumeration of the microorganism Cryptosporidium. The second example addresses the application of MV to the task of intruder monitoring within the context of visual security systems. The chapter will present these two applications to illustrate problems encountered in this type of detection.

In section 2, the chapter presents a general overview of MV applications and discusses some problems associated with the application of MV to objects of high variability. Section 3 will present software architecture of a MV application and its characteristics. Section4 will discuss MV application to variable objects. Section 5 concentrates on two problems associated with this type of application: focus control and error due to changes in illumination conditions. In section 6 AI implementation is discussed and in section 7 Two MV application examples of detection of highly variable objects as are presented.

2. An overview of MV applications

Computer-based vision and automation tools are used in a wide variety of industrial and scientific applications, including electronics, automotive, semiconductor, pharmaceutical, and research applications. These systems perform process monitoring, information gathering, and "on-the-fly" feedback/control to correct manufacturing problems. Research and development into machine vision can be traced back for more than 30 years. The impact of new technologies in machine vision, as well as the historical evolution of this concept can be extracted from published papers in journal, conferences and industrial applications. This field is highly explored at the moment. Table 1 include a list where the generic terms identifying MV applications of more than 10 applications have been implemented; there is a vast potential for applications in variety of topics. A search in BIDS (BIDS 2011) database, using "MACHINE VISION" as keywords revealed 73,773 published articles covering the period 1952 to 2011. Of those 73,773 publications found not all present relevance to application of MV and most of them cover image processing in general term.

Acoustics	Food science	Oceanography	Robotics	Dermatology
Agriculture	Forestry	Oncology	Sociology	Education
Anatomy	Internal medicine	Ophthalmology	Spectroscopy	Engineering
Astronomy	Genetics	Optics	Surgery	Media studies
Biochemistry	Geography	Pathology	Telecommunications	Gastroenterology
Business	Geology	Pharmacy	Toxicology	Geochemistry
Cardiology	Gerontology	Physiology	Transportation	Health care
Chemistry	Immunology	Plant science	Veterinary	Photography
Communication	Infections	Polymers	Zoology	Material science
Biology	Instrumentation	Psychology	Control systems	Mathematics
Education	Mechanics	Nuclear medicine	Behavioural science	Microscopy
Electrochemistry	Metallurgy	Rehabilitation	Biophysics	Management
Energy & fuels	Mining	Remote sensing	Microbiology	Evolutionary biology
Entomology	Neurology	Reproduction	Economics	Surgery
environment	dietetics	Respiratory system	Cell biology	Library science

Table 1. Subject areas with more than 10 MV applications

The early papers refer in general to the idea of image processing and pattern recognition in general terms. A more refined search in BIDS database, using "MACHINE VISION APPLICATIONS" as keywords revealed 1,042 published articles covering the period 1984 to 2010. Figure 1 shows a graph of the number of applications over a 5 years period between 1980 and 2010.

The graph shows a steady increase in the number of applications. In the first 5 years there was an increase from 12 applications in late 1980's to 83 applications in the early 1990's indicating a 7 fold increase. Since its introduction MV has seen an explosion in the number of applications. In the first 10 years the number of application had increased of about 30 times. In general the steady increase in the number of application is due to the increase in computing power, new and better image processing algorithms, better quality in acquisition

of images (hardware) and reliability of Artificial Intelligent (AI) tools. To illustrate the problems of the detection of living/variable objects from the search in BIDS database the articles were classified according to the type of application algorithms/technology. This classification is presented in table 2. By sorting these articles according to the different algorithms/technology used, a general impression on the technology achievements can be approximated. This also provides information of expected development on the MV application.
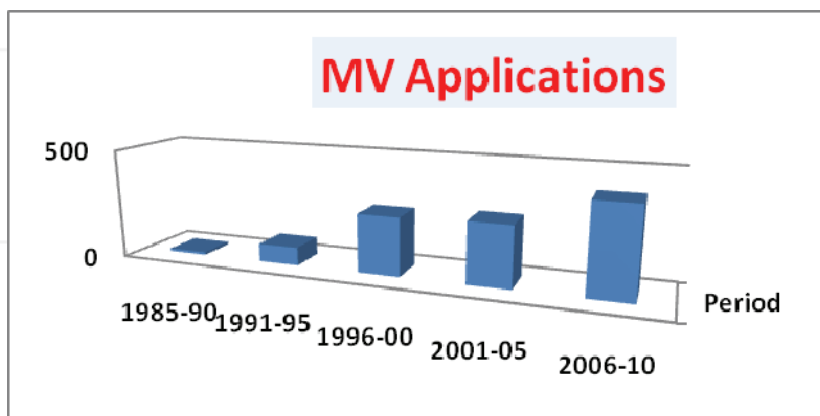

Fig. 1. Number of MV applications between 1980 and 2010.

Field	MV Applications	%
General (theoretical) algorithms or technologies	158	15.16%
Industry related applications	259	24.85%
Control, Instrumentation	124	11.90%
Optics, Robotics, computer science	159	15.26%
Microorganisms, Living cells, variable shape objects	54	5.18%
Agriculture, Food Industry	135	12.96%
Flight Control, Car Control, Automatic Tracking, Traffic Safety	87	8.35%
Textile, Leather, Jewellery	66	6.33%

Table 2. Research Statistics MV applications period 1980-2010

The vast majority of the research and application in the field of Machine Vision is found in the development of general algorithms. The needs of industry strongly influence the type of algorithms developed. The industry generally demands algorithms specialised in pattern recognition, especially algorithms tolerant to lighting variance and partial (dust) occlusion, and occasionally to changes in size (scaling). The industry's demand for gauging application encourages also the development of sophisticated edge detection algorithms, using sub-pixel accuracy (Hanks, 1998). This is also the main reason why powerful Vision Software development platforms such as IMAQ from National Instruments, Visilog from Norpix or PatMAX from Cognex Corporation, (these being just a few examples) appear on the market. The main benefit that these development environments have provided is that it is no longer

necessary to spend long time developing routine algorithms for image processing. Algorithms such as threshold, image manipulation, spatial filtering, Binary Large Object (BLOB) analysing, edge detection, etc. are ready available.   A vision engineer can concentrate on the development of the specific algorithms for each application. Also the management of Operating System related tasks are easier to handle and thus time saving. The movement of computer vision from the "experimental technology" category into industry-strength mainstream applications provides another benefit: the amount of research undertaken in the development of more powerful equipment for image acquisition is increasing, colour cameras with mega-pixel resolutions being widely available this days. The lighting system for machine vision is continuously evolving, different solutions being available with respect to the application demand (Braggins, 2000) e.g. Reflective Surfaces, Undulating Surfaces, Moving Parts. The fourth place in Machine Vision development of algorithms is taken by the agriculture and food industry (13%). The vast majority of these algorithms involve particle sorting e.g. Grain, Olives, Apples, etc or quality tests. This sorting is usually done using BLOB analysis, but advanced algorithms such as combining specific morphological and colour characteristics (Luo et al., 1999) can be found in this field and used for other applications. The algorithms involved in Flight Control, Car Control, Automatic Tracking, and Traffic Safety are usually highly specialized for these fields and are appearing with more frequency. This also applies for algorithms involving texture analysing, developed for textile and leather industry.

The research undertaken in the field of living cells/variable objects represents a total of only 5.18% of the total machine vision applications. The main reason for this low number of applications is the fact that the living cells are not size and shape invariant. On the contrary, the size, shape and sometimes even the colour vary during their life cycle. Hence a really large number of factors have to be taken into consideration when designing a machine vision application in this field. Also in terms of automatic classification the amount of artificial intelligence embedded in these applications is relatively complex and high. The "If-Then, "True-False" logic is usually not suitable, and a different, more advanced approach is needed e.g. Fuzzy Logic, Neural Networks. The intelligence embedded in these applications is relatively high; however the results can be impressive. The application of machine vision to detect this type of objects presents the problems associated with the physical variance of living cells and requires the development of specialized algorithms.

3. Machine vision software architecture

A proposed architecture may include a block to control the camera/microscope and a block to control the image acquisition required to control the hardware. Ideally all controls should be done automatically. The system requires an image processing block with the required image processing algorithms to manipulate and analyse images according to the application. The algorithms will extract information required to provide an output for decision from the system. In many cases the MV should take a decision without human intervention but with the aid of an AI application section. The AI block will add computer capability to the MV system that exceeds human performance in some cases. A MV's software architecture should include all the elements presented in block diagram in figure 2. The MV should provide a way of informing users of the results. The communications block allows information to be provided in various forms/formats including remotely to any place required.
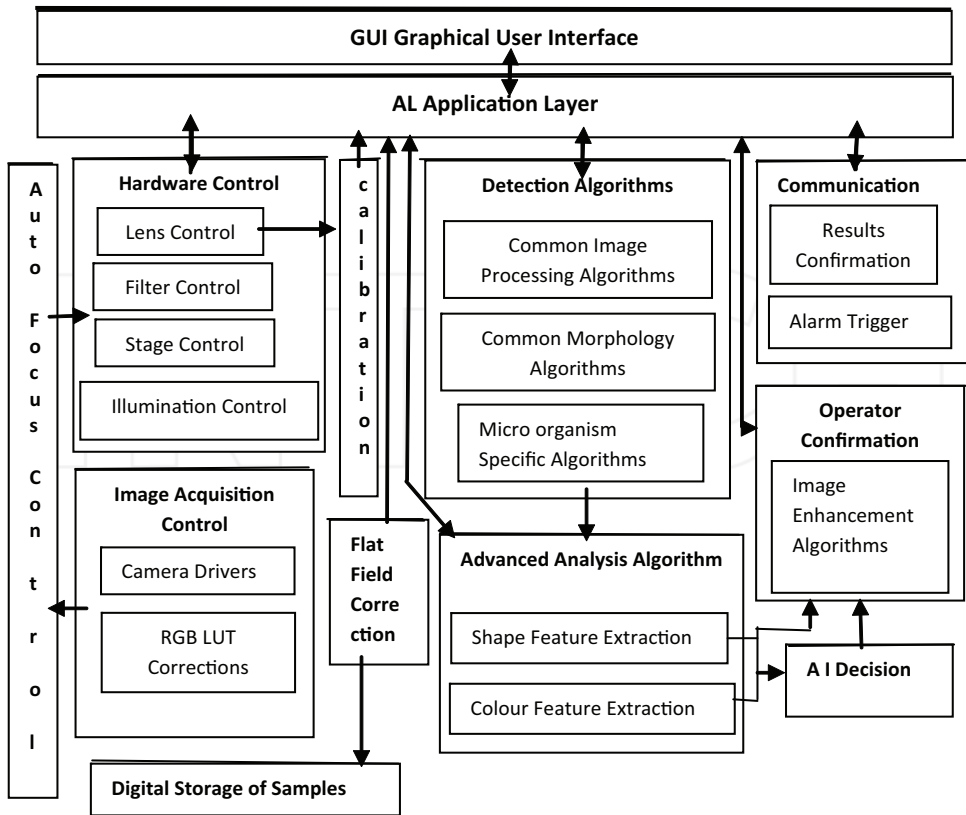
Fig. 2. Machine Vision Software Architecture

All blocks are accessed and organised by the Application Layer (AL), the MV incorporates a Graphical User Interface, providing access to an operator or expert to analyze/confirm results if required.

The MV system presented in figure 2 includes all the features and control algorithms required by a MV application. The AI decision section is tailored to a particular application.

4. MV application to variable objects

To establish a MV application to variable objects it is necessary to consider the general characteristics of the environment in which the MV will operate.

- in which hierarchy the MV will be used (decision making, alarm, identification…).
- application type. It can range from very complex systems with a lot of different information required to be extracted from images to simple applications with easy application.
- frequency of application (mass samples, batch or single analysis).
- level of MV automation (manual, semi automatic, automatic, mixed system).

Depending on the application an MV topology, architecture and algorithms can be planned. Subsequently the relevant outputs variables required for the AI application should be selected. It requires experience to find the best combination of outputs variables. Usually it is required to work with several algorithms at the same time in order to select the ones that perform best. Once the appropriate algorithms are selected they may require to be optimised to increase efficiency of the application.

5. Problems associated with automatic detection of living 'objects'

For both image analysis and image processing, attention has to be paid to errors occurring during the image formation and acquisition, as it is far more difficult if not impossible to eliminate these errors later using image post processing algorithms. There are limitations with respect to the accuracy of the object representation by an acquired image. Both the optics which forms the image of the object on the sensor and the digitisation process introduce errors (Ellenberger & Young, 2000). There are several phenomena which directly influence the acquired image therefore awareness is necessary (Ellenberger & Young, 2000). Living objects have high variability over their life cycle. They can change shape, size, colour or biological structure. In order to detect these objects advance algorithms need to be applied. The success of the advance algorithms application will depend on the elimination or reduction of variability of image due to the acquisition of frames. In this context the variation of image due to change in illumination and change of focus needs to be reduced. Some general problems in image processing affect the detection of objects. In the case of living cells, extraction of information to be input to artificial intelligence for the analysis requires minimizing these problems, in particular the problem with autofocus and problem with illumination. Some solutions are presented in this chapter.

5.1 Problems associated with autofocus control

In order for any image processing algorithms to be effective, the quality of the acquired image must be optimal. In microscopy one of the most important factors is the focus quality, as the even smallest focus deviation produces a blur effect on the acquired image. Therefore the focus control must be automated and performed as often as possible. Under these circumstances an autofocus algorithm must be employed. The autofocus algorithm must be relatively fast, involving the minimum computational effort possible. The image is brought to focus on the camera controlling the tuning of focus lenses stage using a stage controller. Auto-focus could be achieved by developing a feedback focus controller as exemplified in figure 3. This is possible if a measure of the focus quality in a numeric form is extracted (focus score). The hill climbing algorithm (Russell & Norvig, 2002) can be iterated with focus scores until the best focus is found.

The focus quality of an image can be linked with the perceived "sharpness" of the acquired image, therefore the number and the strength of edges in it. The number of edges can be determined using an edge-extraction algorithm. One of the fastest algorithms to extract edges in any direction is the Laplacian Edge Detection as it involves only one convolution. One problem associated with edge detection algorithms is the noise influence; therefore noise reduction algorithms are necessary. A good choice could be a median filter, as it has a minimal influence over continuous edges.
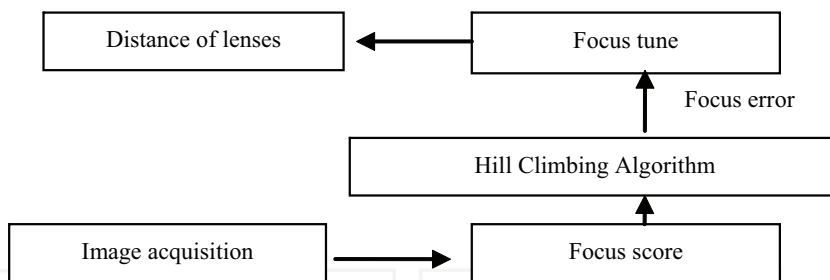
```
┌─────────────────────┐        ┌─────────────────────┐
│  Distance of lenses │ ◄───── │     Focus tune      │
└─────────────────────┘        └─────────────────────┘
                                         ▲
                                         │  Focus error
                               ┌─────────────────────┐
                               │ Hill Climbing Algorithm │
                               └─────────────────────┘
                                         ▲
                                         │
┌─────────────────────┐        ┌─────────────────────┐
│  Image acquisition  │ ─────► │     Focus score     │
└─────────────────────┘        └─────────────────────┘
```

Fig. 3. Autofocus control

5.2 Autofocus control implementation

A series of images of a test pattern are acquired using the minimal resolution possible and greyscale representation to minimise the computational effort. These images are shown in figure 4, in the acquisition order, with the first 2 being completely out of focus, and the other being close to optimal focus:

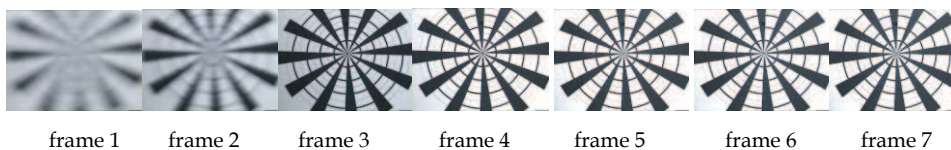| frame 1 | frame 2 | frame 3 | frame 4 | frame 5 | frame 6 | frame 7 |

Fig. 4. Focus test frame 1 to 7

From each image the light plane is extracted and a Laplacian Edge Extraction algorithm is applied. The kernel used in this case was

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}. \tag{1}$$

One of the resulting images is presented in figure 5(a).

(a) Test frame Edge information   (b) Median filter applied

Fig. 5. Edge extraction

A median filter defined on a small neighbourhood is applied to minimise the noise effect on the final focus score. The result is shown in figure 5(b). Although the extracted edges are slightly affected, all continuous edges are still present while most of the background noise

has been eliminated. A larger neighbourhood may be considered to improve accuracy, but a significant increase in computational time has to be expected. A sum of all pixel values is then performed, the result being a single numerical value. By applying this algorithm on 10 test images the results presented in Table 3 are obtained. By comparing the focus score with the test images, it can be seen that they are closely tied up.

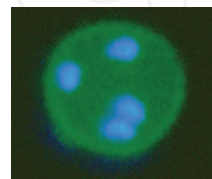| Image	Focus Score	Image	Focus Score
Frame 1	1317531	Frame 6	5765586
Frame 2	1374433	Frame 7	5873043
Frame 3	5374843	Frame 8	5879743
Frame 4	5902432	Frame 9	5868048
Frame 5	5756735	Frame 10	5868667

Table 3. Focus Measure Results

On the first two images, the score is relatively low, as they are completely out of focus. On the last seven the focus score is comparable, with the 4th image obtaining the best result. When comparing some of these frames the human eye will have difficulties detecting the optimal focus. The focus score will differentiate the best focus in frames that appear to have the same focus. The algorithms presented here have the advantage of high processing speed in reaching the best focus, in the order of milliseconds in an average PC system. Real time software controlled autofocus is possible.

5.3 Problems associated with illumination

It is possible to combine several light sources that illuminate the specimen in different ways. The monochromatic or narrow band light sources emit light at a single wavelength or a very narrow band (e.g. Laser, which is also coherent). Halogen and mercury lamps are broad band incoherent light sources. Such an incoherent light source can be converted to quasi coherent by closing the aperture to a pinhole (Pawley, 1995). Acquiring images using coherent illumination has some advantages and also some disadvantages: a sharp edge will show ringing effects and the edge will be shifted into the bright area. Also the images look granular (speckle effect), and any aberration or dust on the lens can produce disturbance in the image. However, the most important advantage is that resolution achievable can be better than with incoherent illumination (Goodman, 1996). One of the most important advantages of incoherent illumination is that the image brightness is not changed with modifications of the focal position, a very important consideration when auto-focus algorithms are employed. Another significant aspect for a Machine Vision System is the Image Representation. Colour images can be represented in several different formats (also called Colour Spaces).
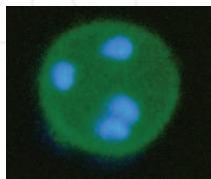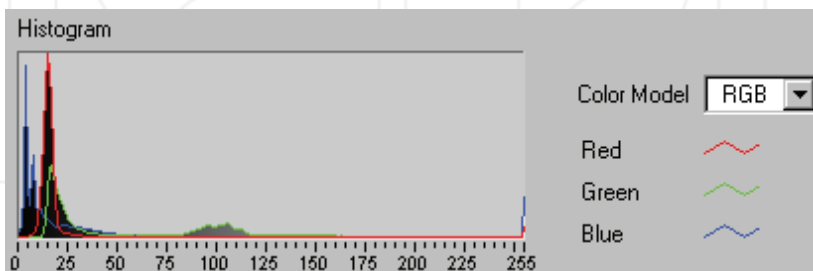
(a) Oocyst under normal illumination   (b) Same image 10% brighter

Fig. 6. Cryptosporidium oocyst with 10% change in illumination.

The most common representation of colours is the RGB colour space, as most image sensors provide data according to this model. RGB colour plane is suitable for image capture and reproduction, but can be very inconvenient for feature extraction. Problems are noticed when minor changes in the lighting conditions occur. Figure 6 shows a sample of a microorganism oocyst common in drinking water known as cryptosporidium. In figure 6 the illumination has been increased by 10% in capturing the image of this cryptosporidium oocyst. Figure 7 presents the histogram of the same cryptosporidium oocyst image with 10% difference in lighting conditions. The histograms are very different and make it difficult to separate the object of interest and its background.

(a) Image normal illumination
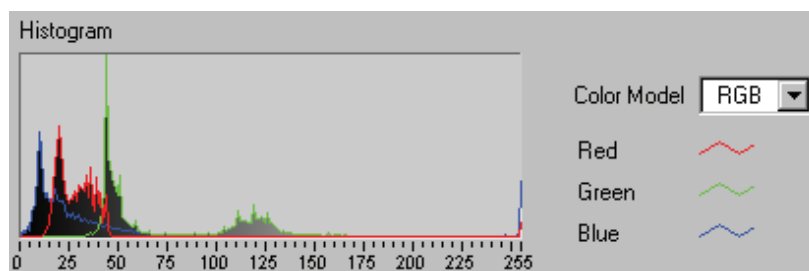
(b) Same image with 10% change in illumination

Fig. 7. RGB histograms

One solution to overcome this problem is the use of a linear transform from the RGB colour space into an alternative model such as HSL – Hue, Saturation and Lightness, or HSI – Hue, Saturation and Intensity. Hue refers to the perceived colour (viewed technically as the dominant wavelength), saturation refers to the dilution of the colour by white light, and the lightness (or value) refers to the sensation of brightness relative to a white reference. The advantages of HSL over the RGB colour space are illustrated in figure 8 where the same 10% change in lighting condition is applied to the same cryptosporidium oocyst object of the previous histograms.

The results of the histograms indicate that the Saturation plane is slightly affected and the Hue plane is relatively unaffected. The HSL space will make it possible to isolate the object of interest from the background. The RGB-HSL conversions are computational intensive and hardware conversions are preferred.
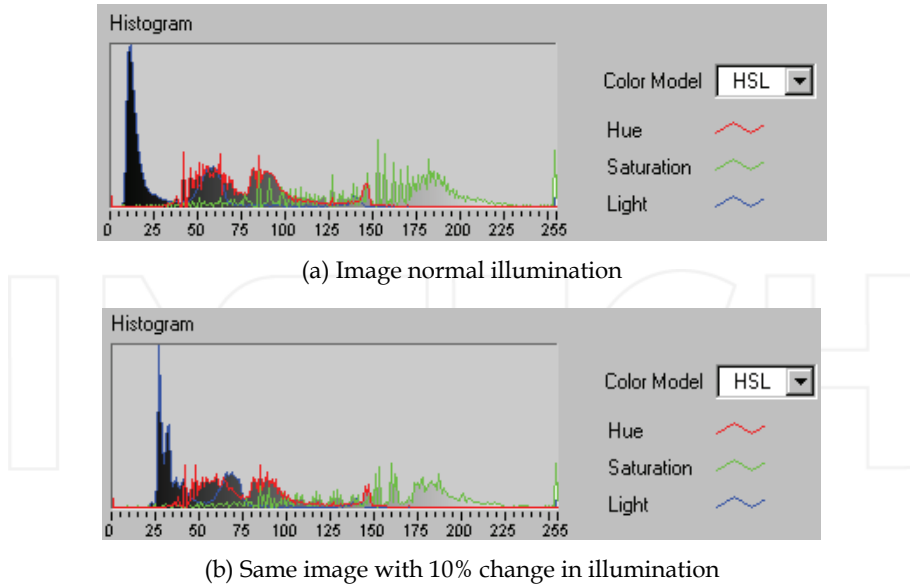
(a) Image normal illumination

(b) Same image with 10% change in illumination

Fig. 8. HSL histograms

6. Artificial intelligence applied to MV

In many cases the MV should take a decision without human intervention but with the aid of an artificial intelligent (AI) application section. The AI application will depend on:

- the type of application.
- the outputs from the vision algorithms.
- the testing/training samples available.
- availability of training vectors statically independent.

Computers already emulate some of the simpler activities of the human mind. They can perform mathematical calculations, manipulate numbers and letters, make simple decisions, and perform various storage and retrieval functions. In these applications, the computer is exceptional and usually exceeds the human mind in performance. Artificial intelligence gives computers added computing capability, allowing them to exhibit more intelligent behaviour. The most common types of artificial intelligence system are Expert Systems, Fuzzy Logic and Artificial Neural Networks (ANN). Providing a set of parameters that would completely and undoubtedly describe highly variable objects is extremely difficult. Choosing to implement the classification process by means of an ANN seems optimal, as all it should achieve is to provide a suitable training set to a carefully chosen neural network. Because of the variability of the objects in some cases problems are encountered with this type of AI in that there are not sufficient numbers of training samples to train the ANN. Under this circumstance confident training cannot be performed successfully. For this type of application a more flexible approach is needed to take into account different variation of the objects. Fuzzy logic system can in some cases successfully address this issue, as it logically implements degrees of membership based on likelihood. Therefore a fuzzy logic inference engine is in many cases the preferred AI decision tool in this type of applications.

7. Examples of application to variables objects

Two examples of MV application to objects that are highly variable are presented in this section. The first example detects the existence of Cryptosporidium in water and the second example detects suspicious activity of humans using Closed Circuit Television (CCTV).

7.1 Example 1. detection of cryptosporidium in water

Cryptosporidium has been widely recognized as a serious cause for concern, with a very large number of waterborne infections caused by its oocysts. In its transmissive stage – the oocyst - is a frequent inhabitant of raw water sources used for the abstraction of potable water. Its importance is heightened because, coupled with its low infection dose, conventional water treatment process, including chemical disinfection, cannot guarantee to remove or destroy oocysts completely. Waterborne transmission is well documented (Smith & Rose, 1990, 1998), and can affect a large number of individuals. More than an estimated 427,000 have been affected in 19 documented waterborne outbreaks (Smith & Rose, 1998). Transmission is via an environmentally robust oocyst excreted in the faeces of the infected host (Smith & Rose, 1998). At least 21 species of Cryptosporidium have been named (Frazen & Muller, 1999). Cryptosporidium parvum is the major species responsible for clinical disease in human and domestic animals (Current, 1988; Currents & Gracia, 1991).The laboratory diagnosis of Cryptosporidium is dependent upon demonstrating oocysts in the sample by microscopy. Here, oocysts must be distinguished from other, similarly shaped, contaminating bodies present in the sample, and the microscopic identification of oocysts is dependent upon morphometry (the accurate measurement of size) and morphology. The current manual is expensive, labour intensive, time consuming and unreliable. Cryptosporidium is difficult to identify because of its size and morphology. Regulatory bodies from all over the world acknowledge the continuous monitoring of water sources for Cryptosporidium as imperative. Many requirements, rules and regulations are in place to attempt to address the control of Cryptosporidium which threatens the safety of drinking water. As an example the EEC produced a drinking water directive (EEC, 1998). The example presented here complies with this directive. The cryptosporidium structure in a computer generated model is shown in figure 9. Samples are stained and dried onto microscope slides and micro-organisms are detected using fluorescence microscopy. The drying process causes oocysts to collapse which in some cases leads to shape distortion and the physical release of sporozoites. This made the manual detection difficult.
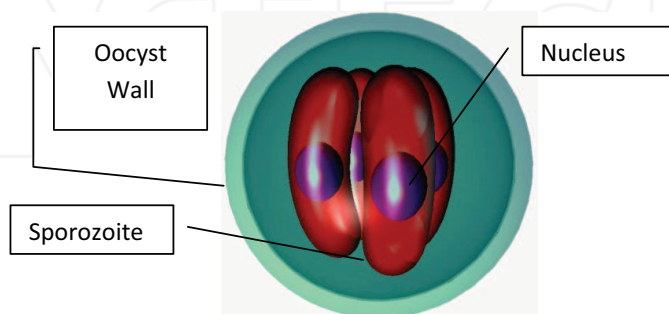

Fig. 9. Computer generated model of Cryptosporidium

The acquired image of a slide containing Cryptosporidium, viewed under Differential Interference Contrast (DIC) is presented in figure 10. Figure 10 shows a sample containing cryptosporidium oocysts.
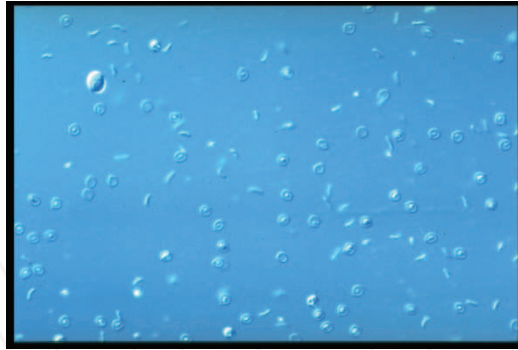
Fig. 10. Water sample containing cryptosporidium oocysts

Significant contrast is achieved using a highlight details filter and elimination or reduction of noise. The binary noise present is efficiently eliminated using the following algorithm: (Fernandez-Canque et al, 2008, 2000).

i.   A buffer copy of the image to be cleaned up is generated.
ii. Two successive erode functions are applied on the original image.
iii. All pixels from the buffer copy 8-connected to the non-zero pixels from the image are added to the image
iv. Step (iii) is repeated until no pixel is added.

Figure 11 (a) shows part of the sample with the noise eliminated and with the missing pixel reinserted.

(a)   (b)
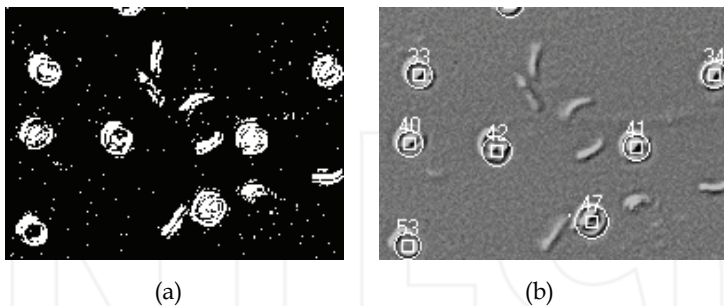
Fig. 11. (a) Sample with threshold images combined (b) Sample final result oocysts identified

The next step is to reinsert the missing pixels within the objects boundaries (Fernandez-Canque et al, 2000). A closing algorithm is performed, using a kernel size of 5. A NOT function is performed, followed by a labelling function. The result is an image which has a value of 0 associated with all objects, a value of 1 associated with the background and a value greater than 1 for every hole in the objects. By replacing the values greater than 1 with

0 and negating the image again we achieve holes filling. All objects too small to be a Cryptosporidium are eliminated. This is achieved using the same algorithm as for binary noise removal, but with 7 erosion functions applied. Then a distance function is applied, and a circle detection algorithm (Danielsson, 1980) is used for Cryptosporidium detection. The result is presented in Figure 11(b).An advance analysis algorithm can use different colour planes to allow the determination the existence of the 4 nucleus in an oocyst. On the green plane, both the wall and the nucleons of the Cryptosporidium are detected. Figure 12(a) show the wall and nucleons of a cryptosporidium oocyst after noise is reduced and objects of interest are separated from background.

(a) Green Plane wall and nucleons   (b) Red plane nucleons

(c) Red Plane noise reduced   (d) Wall: subtraction of green and red plane
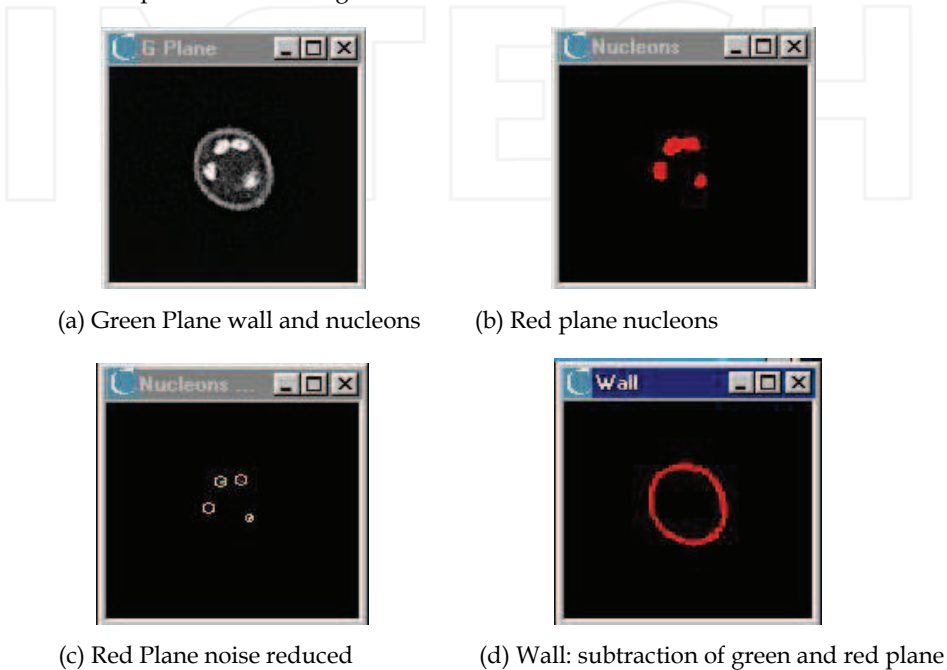
Fig. 12. Green and Red plane

On the red plane, only the nucleons are detected. This gave an easy method of separation. Figure 12(b) shows the red plane after the image is cleaned of noise. On the red plane a look-up table function was applied. Finally, threshold and binarisation was used, followed by noise removal and BLOB closing. Figure 12(c) shows clearly the 4 nucleons to identify cryptosporidium. By subtracting the nucleons from the green plane in binary format the wall was obtained and measurements were done, this is shown in figure 12(d). The morphological characteristics of cryptosporidium can be extracted unequivocally. Image processing allows the manipulation of these figures at the pixel level to analyse details of very small dimensions. As Cryptosporidium have a diameter in the range of 4 to 6 microns, a human operator would find difficult to identify this micro-organism under the microscope. This provides a method of identification of cryptosporidium. After the completion of the Advanced Analysis the system goes into the next stage – A I Decision Making. Expert Systems have problems with flexibility for this application and ANN

encounter problems with training. It was found that the use of Fuzzy Logic suits this application by mimicking human knowledge based decision making. The classification is based on the features extracted by the Advanced Analysis Algorithm and a customisable rule base. The proposed approach allows a reliable detection of waterborne micro-organisms in large quantities of water and outperforms the current manual detection in term of cost, time of results, accuracy and reliability (Fernandez-Canque et al., 2009).

7.2 Example 2: Detection of suspicious activity

In recent years, the use of surveillance cameras has increased in popularity. This is partially due to reduction in cost and technological advances. CCTV systems have become very popular in observing public places. Current technology makes provision for an operator to examine live surveillance footage from remote locations as they can be transmitted over the internet, cables or wireless mediums. In this example the MV application detects suspicious activity automatically by studying human posture and observing full trajectories of people (Fernandez-Canque et al., 2009). In this study, work has been carried out with the aim of achieving fully automatic detection of intruders using a static camera and in real time. CCTV has the advantage that relatively large areas can be monitored and intruders can be seen as compared to other detection methods. The main use of CCTV is based on reaction to a past incident by revising image recorded; the aim of this work is to make the use of CCTV more efficient by assessing suspicious activity in an active manner and alert operators to an intrusion. By achieving an automatic detection some problems associated with this type of surveillance can be avoided. It is known that the span of concentration of any operator is very short (Saarinen & Julesz, 1991; Goolkasian, 1991), and there is unreliability due to operator's fatigue and poor detection due to large number of irrelevant images known as Eriksen effect(Eriksen & Murphy, 1982; Eriksen & Hoffman, 1972).

7.2.1 Intruder detection

To detect an intruder, the background with no intruders present is recorded while the camera transmits the video to the computer, and each frame is analyzed. After reduction of noise and distortion, frames are compared to the original image of the plain background with no intruder present. This process results in accruing the pixels of the object of interest. Pixels in this object are counted, large numbers these pixels reflect a significant background change, thus an intruder is likely present. If no intruder is detected, the background image is replaced with the current frame of the video, which is compared against the next video frame as the program repeats this cycle once. A pixel count is then applied to confirm if an object has been detected, and finally the parameters can be computed for a bounding box. This sequence of steps will encapsulate any detected objects in its bounding box, maintaining the original grey-levels of the object. This sequence of actions is illustrated as block diagram in figure 13.

The bounding box contains the coordinates of the intruder. This object of interest within this bounding box can be used for further analysis in the detection of suspicious activity. This simple process will allow the detection of an intruder. The comparison of frames can be efficiently implemented after noise is eliminate or reduced as indicated in the previous section.
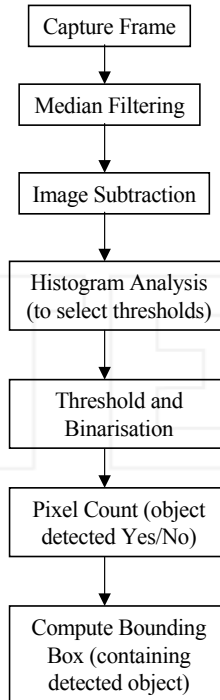
```
┌─────────────────┐
│  Capture Frame  │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Median Filtering│
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Image Subtraction│
└─────────────────┘
         │
         ▼
┌─────────────────────┐
│  Histogram Analysis │
│ (to select thresholds)│
└─────────────────────┘
         │
         ▼
┌─────────────────┐
│   Threshold and │
│   Binarisation  │
└─────────────────┘
         │
         ▼
┌─────────────────┐
│ Pixel Count (object│
│ detected Yes/No) │
└─────────────────┘
         │
         ▼
┌─────────────────────┐
│ Compute Bounding    │
│ Box (containing     │
│ detected object)    │
└─────────────────────┘
```

Fig. 13. Block diagram object detection.

7.2.2 Motion analysis

A particle analysis is performed on the binary image produced after thresholding. This analysis calculates two measures for each detected object or pixel cluster within the binary image. A pixel cluster is a grouping of connected pixels. The measures calculated are *pixel area (size)* and *shape factor*.

Pixel area is the number of pixels of intensity value 1 (white) in a given cluster. X is the cluster containing all pixels in the object of interest. The *area* of an object can be calculated as a relative measure with respect to the entire image area. In the discrete case, it is approximated by the number of pixels in the object of interest.

$$A(X) = \sum_{i,j} g(x_i, y_j) \tag{2}$$

Where $A(X)$ is the area of the object X, $g(x_i, y_j) = 1$ if the pixel lies within the object X *and* $g(x_i, y_j) = 0$ otherwise

The *area* measure can be used given a binary image to obtain the relative size of the object. This measure can give an indication of whether the object is sufficiently large (in relative terms) to warrant further analysis. It can also be weighted by a size factor which characterises the area of a single pixel. In that case, the measure is physically homogeneous

to an area. A shape cannot be characterised by a single measure, or even a simple set of measures. Several quantities have been defined to account for specific shape characteristics.

The *shape factor*, F_C, is defined as:-

$$F_C = \frac{L(X)^2}{4\pi.A(X)} \tag{3}$$

Where $A(X)$ is the area of the object X *and* $L(X)$ is the perimeter of the object, defined below.

The perimeter designates the length of the object boundary. In the discrete case, the perimeter can be simply estimated as the number of points which lie on the object boundary. The *shape factor* measure is invariant to rotation, reflection and scaling. It has no dimension and is equal to 1 for a disk. It measures the elongation of an object. An elongated set has a higher *shape factor*. This measure, with its ability to distinguish the elongated form of a human figure, is particularly useful in the object detection process. The *area* and *shape factor* for all pixel clusters (which are potential objects of interest) within the binary image is calculated. A fixed threshold for each measure is used to select objects for further analysis. Any small differences in pixel intensity between the captured frame and the previously acquired frame will be effectively removed during the *area* and *shape factor* thresholding operation. Small differences do not indicate the presence of an object in the frame, but can often be attributed to small changes in illumination between successive frame captures. The threshold chosen for each measure is scene dependent. Once the approximate size of objects of interest is known relative to the size of the entire image, thresholds can be calculated to ensure only object *areas* greater than the threshold are retained for further analysis. After thresholding, the resulting binary image will contain the pixel clusters corresponding to objects meeting the selection criteria. The object detection algorithm is performed on each frame acquisition. Detected objects that have met the size, shape and Si factor (as defined in equation (7) Fernandez-Canque et al, 2009) criteria are tracked in terms of their position within camera images. The tracking algorithms require each object to be represented by a single pixel co-ordinate position within the image. The *barycentre* of detected objects is used to provide the single co-ordinate. The *barycentre* of an object is similar to its centre of gravity. The resultant co-ordinate position for each object gives a uniform positional locator between successive frame captures which is independent of object size and shape. The inertia moments define some global characteristics of the object but it is the first order moments of inertia which define the *barycentre*. They are defined in the discrete case as:-

$$M_{1x} = \frac{1}{A(X)} \sum_X x_i \tag{4}$$

and

$$M_{1y} = \frac{1}{A(X)} \sum_X y_j \tag{5}$$

Where M_{1x} is the first moment of inertia in the *x* plane.
            M_{1y} is the first moment of inertia in the *y* plane
            $A(X)$ is the area of the object X
            (x_i, y_j) is a point in the object

The *barycentre* of each detected object is passed to the tracking algorithm after every frame acquisition. The positional locator of each object is passed to the tracking algorithm. Subsequent frame acquisitions provide a new positional locator for each detected object. The tracking algorithm computes the linear distance from every initially detected object to every object detected in the subsequent frame acquisition. The shortest distance between each initially detected object and subsequently detected objects is selected and the object which lies the shortest distance from the initial object is then determined to be the same object as in the previous frame. The process is repeated for each frame acquisition thus allowing objects to be tracked. The distance, *L,* between 2 co-ordinate positions is calculated as follows:-

$$L = \sqrt{(X_1 - X_0)^2 + (Y_1 - Y_0)^2} \tag{6}$$

Where (X_0, Y_0) is the co-ordinate position of the first point

(X_1, Y_1) is the co-ordinate position of the second point

Objects detected between consecutive frames are selected for *Si factor* analysis. Each object pair cannot always be assumed to be the same object between consecutive frame acquisitions. The *Si factor* provides one method for determining that tracked objects between successive frame captures are the same object within the images. The Si factor can be calculated as follows:-

$$Si = \frac{\frac{\left|(A(X_{nI1}) - (A(X_{nI2}) -\right|}{A(X_{nI1})} \times 100 + \frac{\left|(F_c(X_{nI1}) - (F_c(X_{nI2})\right|}{F_c(X_{nI1})} \times 100}{2} \tag{7}$$

Where $A(X_{nI1})$ is the *area* of object X_n in Image 1

$A(X_{nI2})$ is the *area* of object X_n in Image 2

$F_c(X_{nI1})$ is the *shape factor* of object X_n in Image 1

$F_c(X_{nI2})$ is the *shape factor* of object X_n in Image 2

The *Si factor* is calculated for all *n* objects detected and provides a confidence measure to determine that objects tracked between images are the same object. The lower the Si factor, the more the object detected in the subsequently acquired frame conforms to the size and shape characteristics of the object in the previously acquired frame. Thresholds may be set for the allowable values of Si factor. The value of such thresholds will vary depending on the scene being viewed and the natural variations of objects in the same class that are to be detected. Objects detected between successive frames that have a Si factor which lies above the threshold can be assumed to be different objects. This provides the capability for the tracking algorithm to detect when an object has been lost rather than tracking the incorrect object. The Barycentre of each detected object is passed to the tracking algorithm after every frame acquisition. Subsequent frame acquisitions provide a new position for each detected object. The tracking algorithm computes the linear distance from every initially detected object to every object detected in the subsequent frame acquisition. For the external scene, a sufficiently large area is covered by the camera to allow object tracking to be verified. The

object detection algorithm is applied to capture sequential frames to determine initially if an object has been detected, then to track any such object.

```
┌─────────────────────────────────────────────────────┐
│      Pixel clusters obtained from binarised image     │
└─────────────────────────────────────────────────────┘
                           │
┌─────────────────────────────────────────────────────┐
│   Pixel Area and Shape Factor calculated for each cluster │
└─────────────────────────────────────────────────────┘
                           │
┌───────────────────────────────────────────────────────────────┐
│ Threshold applied based on Pixel Area and Shape Factor (yields detected object) │
└───────────────────────────────────────────────────────────────┘
                           │
┌─────────────────────────────────────────────────────┐
│      Barycentre and Si Factor of Object Calculated    │
└─────────────────────────────────────────────────────┘
                           │
┌─────────────────────────────────────────────────────┐
│                  Next frame acquired                  │
└─────────────────────────────────────────────────────┘
                           │
┌───────────────────────────────────────────────────────────────┐
│   Algorithm repeated to obtain Barycentre of next detected object │
└───────────────────────────────────────────────────────────────┘
                           │
┌─────────────────────────────────────────────────────┐
│      Pixel distance between objects obtained          │
└─────────────────────────────────────────────────────┘
                           │
┌─────────────────────────────────────────────────────┐
│          Si Factor of objects calculated              │
└─────────────────────────────────────────────────────┘
                           │
┌───────────────────────────────────────────────────────────────┐
│ Objects with similar Si Factor deemed to be same object between frames │
└───────────────────────────────────────────────────────────────┘
                           │
┌─────────────────────────────────────────────────────┐
│            Process repeated as object tracked         │
└─────────────────────────────────────────────────────┘
```
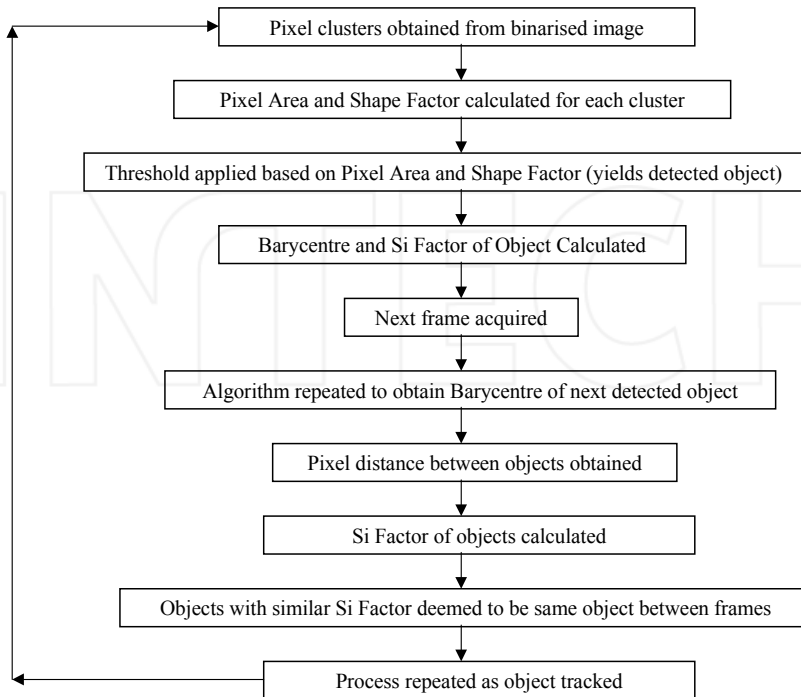
Fig. 14. Block diagram. Object tracking.

Given the small size of objects in the external scene, the additional processing steps are carried out. These steps include the analysis of detected clusters yielding data on object size and shape characteristics. Such analysis is not required for an internal scene because any detected objects would be sufficiently large in comparison to the total area of the scene that errors due to noise could effectively be eliminated by setting large thresholds. However, in the external scene, any detected object will yield a small pixel area so setting high thresholds would likely cause even objects of interest to be thresholded out. The size and shape characteristics can be used to assist in not only detecting objects, but also to subsequently track them. Kalman filter (Kalman, 1960) can provide a good way of tracking an object of interest, in the case of identifying the object the use of the Si factor may be more useful. Especially when objects of similar characteristics are physically close the Si factor can distinguish between objects. A summary in block diagram for the additional series of processing steps for the external scene to track an object of interest are specified in figure 14. The following example details the application of the outlined algorithms to the external scene and also to the task of object tracking.

The captured image in figure 15 is a car park area, and also contained a pedestrian footpath. Figure 15 shows captured frames containing one person.

Fig. 15. Frame car park area containing one object of interest: one person

Figure 16 shows two frames used in the detection of suspicious activity. The example uses a sequence of these frames to track a person in this car park. Cluster will appear for objects of no interest such as birds, movements of branches, etc. After the detection and elimination of unwanted noise, figure 16(c) shows the detected person of the two scenes superimposed.

(a)   (b)   (c)

Fig. 16. Two consecutive scenes (a) and (b) of a captured sequence of frames from a car park scene. In (c) the object of interest is shown for the 2 scenes after detection algorithms have been applied.

For the first frame of figure 16 the object detection algorithm is applied to this image producing the thresholded and binarised image for this frame. This frame contains 44 pixel clusters. The *area* and *shape factor* for each cluster is calculated.

Table 4 shows the analysis of the detected clusters. The minimum and maximum values for *area* and *shape factor* show the range of these measures across the 44 detected clusters.

Threshold selection is determined for a given scene based on the pixel size of the object type to be detected.

This thresholding operation yielded one detected object of *area* 211 (the largest cluster *area*), *shape factor* 3.11 and *barycentre* 102,137.

	Area	Shape Factor	Barycentre
Minimum	1	0.47	n/a
Maximum	211	3.21	n/a
Threshold	150	2.5	n/a
Object	211	3.11	102.137

Table 4. Analysis of Clusters detected

The pixel clusters which have been filtered out are caused by changes in the image between subsequent frame acquisitions. These clusters can be attributed to the movement of the tree branches in the wind. The analysis of each detected cluster and subsequent selection operation successfully removes these unwanted particles to leave only the desired object. The modified object detection algorithm can now be applied to 2 consecutive frame acquisitions. The object tracking algorithm is iteratively applied to each detected object in the initially acquired frame. The first order moments of inertia in the x and y plane are calculated for all detected clusters in each frame acquisition, and used as a positional reference by the tracking algorithm. The completion criteria for this algorithm is met when the nearest neighbouring cluster to each object in the initially acquired frame is found in the subsequently acquired frame. The object detection algorithm, when using the *Si factor*, is more immune to noise.

A sequence of images is captured from an external scene, figure 17 shows a test scene used. To determine suspicious activity the car park image is divided into four distinct areas, as shown in figure 17. The scene has four distinct areas: Area 1 pathways, Area 2 car park, Area 3 exit/entrances, Area 4 perimeter area. Risk Index takes values between 0 and 1, where 0 represent the lowest risk. For this exercise the areas were given the following risk indices: Area 1 risk 0.2, area 2 risk 0.6, Area 3 risk 0.7, and area 4 risk 0.8. The risk indices are determined based on the knowledge of the scene and the threat associated with human movement within that area. In the image presented, the Risk Indices have been chosen such that the higher the Risk Index the greater the risk associated with human movement within that area.

In the experiment conducted, human movement is represented by crosshairs within the segmented images. The crosshairs on each segmented image represents the time-series motion of a human walker in the image. Figure 17 shows the movement of a person within the segmented areas of the scene analysed.

The aim of analysing the time series motion is to elucidate the information relevant to the task of ultimately identifying suspicious activity, and present this data in a format which the

detection system can use. The aim of this experiment is to automatically determine if a given activity, identified by the time series motion of a human walker in conjunction with the segmented image data, can be interpreted as suspicious. The information used by the detection system should incorporate both the walker's positional data, as well as the Risk Index data.
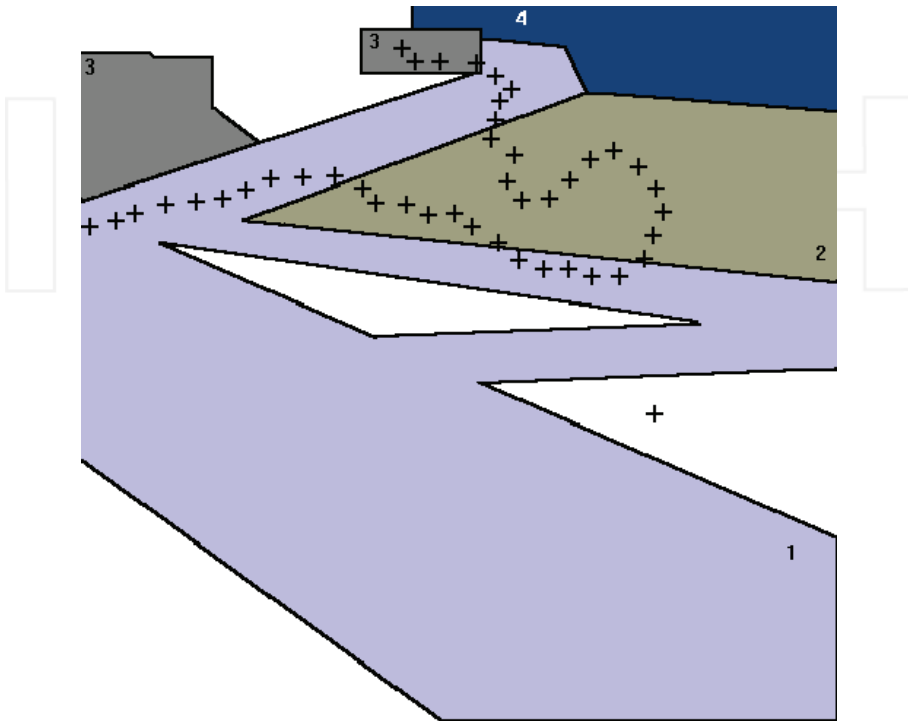

Fig. 17. Trajectory of a human walker within a scene

From this tracking scene information of speed, direction and presence in any zone is extracted and presented to a trained ANN for determination of suspicious activity.

In this experiment, twenty five segmented images are presented, each identifying the path a human walker takes within the scene. Figure 18 shows the response of the ANN to the 25 path samples of people walking patterns in the different areas of the scene analysed indicating the degree of suspicious activity for each pattern.

The ANN results of suspicious activity show a good correlation with the human operator response within scenes. This provides a good forewarning role allowing further investigation by a human operator. The MV application presented here can provide an indication of suspicious activity as the output of the ANN. This ANN response is based on the direction of the intruder, the speed of movement, the position on the different risk areas and the pattern of movement within the scene.
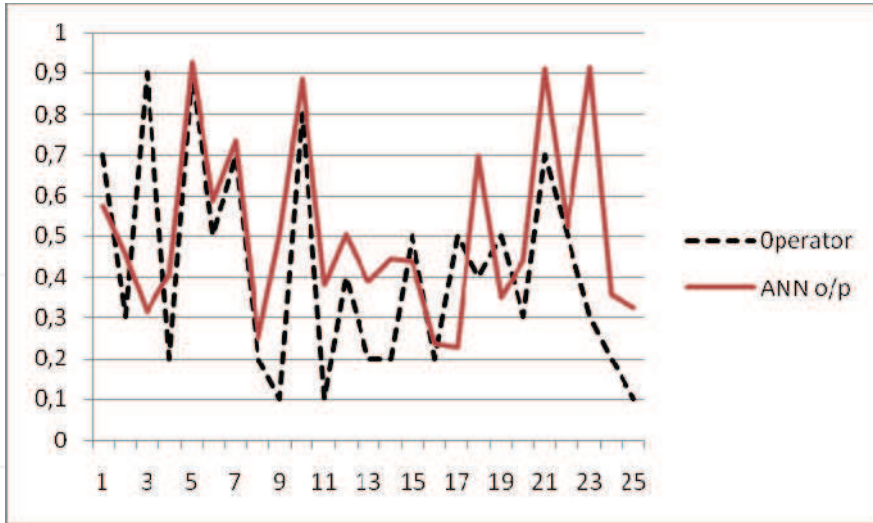
Fig. 18. Response from the ANN and operator to 25 segmented image samples

7.2.3 Posture analysis

 If the person is not moving in areas that are considered a risk, his or her posture may indicate suspicious activity. Each image is subjected to a reduction algorithm, producing a quantised image, followed by a 16 by 16 data array to be presented to the ANN.

The ANN used is trained to provide an output for waling, standing and crouching postures. Figures 19-21 shows results from a trained ANN to a set of 26 images containing humans in these positions.
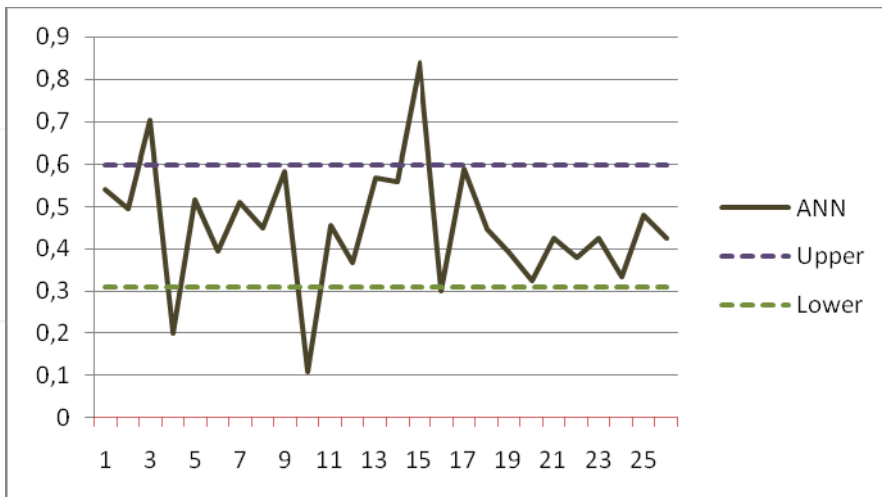
Fig. 19. Classifier results from the ANN for walking postures
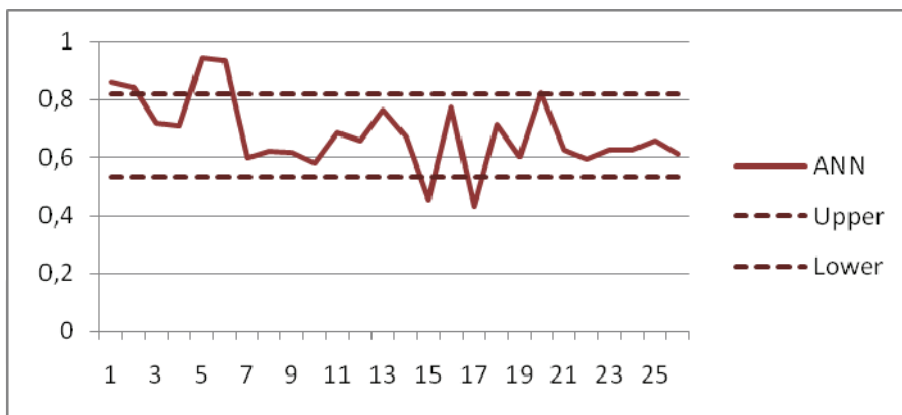
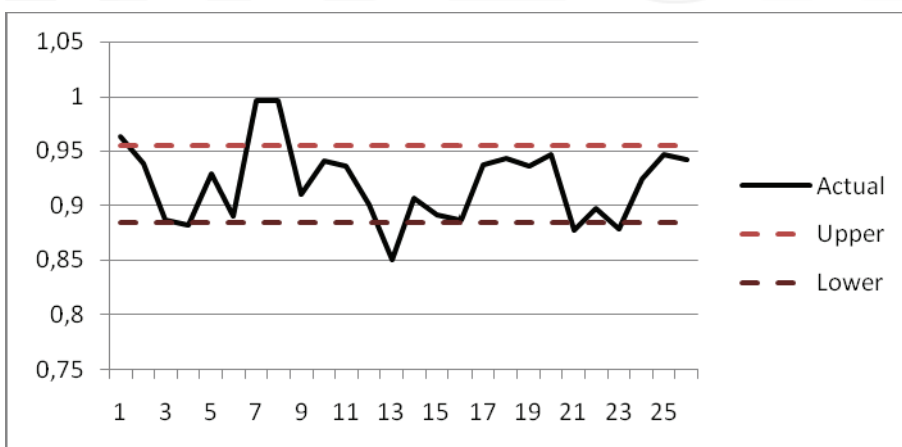Fig. 20. Classifier results from the ANN for standing postures

Fig. 21. Neural Network response for crouching postures.

Figure 19 shows results for walking postures, Figure 20 for standing postures and Figures 21 for crouching postures. The ANN has been trained in such way that values on the y-axis between 0.3 and 0.6 indicate that a walking posture has been detected. Values between 0.52 and 0.81 indicate that a standing posture has been detected and values between 0.88 and 0.95 indicate detection of a crouching posture.

On the basis of the results obtained for each newly analysed posture, upper and lower limits (shown by chain lines) have been set which provide a range of values within which lies a high probability that a given posture has been detected. The walking posture test images included various viewing orientations. In the case of a person walking directly towards or away from the camera, difficulties are encountered in determining if the person is standing or walking. These walking postures are very similar in appearance to standing postures when presented as a two-dimensional camera image. This similarity, and consequent classification difficulty, was predicted for a two-dimensional image system. A number of the

images in the walking posture data set show a higher than expected value although only 15% lie outside the upper and lower limits. The higher values obtained tend towards the values expected for a standing posture. The results produced for the crouching postures show values which are consistently very close to the expected values. The error in the crouching posture data was smaller and more uniform than for the other two posture types. The apparent similarity between a few walking and standing postures caused the classifier some difficulties. A crouching posture is easier to identify from a two-dimensional image. Since crouching could be interpreted as potentially the most suspicious posture, the system developed shows higher performance at detecting the postures of more relevance to the task in hand.

The results are close to the expected response. This can also be used as a warning role as the MV can determine whether suspicious activity is present in an open scene under CCTV surveillance

8. Conclusions

MV application trend indicates that the advance of computation power and hardware developments allows the number and type of application to increase substantially in particular in the last 15 years.

The application of MV to variable objects requires the consideration of various factors to be accurate. Algorithms are required to obtain clear and detailed morphology including advance process using advance algorithms. In this type of application where information obtained in different planes it is useful to eliminate problems produced by variation of illumination. A good focus control based on a focus score allows the MV application to manipulate objects of small dimensions to gather more details of the characteristics of the object to be analysed.

The Machine Vision example presented in this chapter can perform automated analysis to determine whether or not micro-organism oocysts are present in treated water. The system can reliably determine the presence of micro-organisms and enable samples to be accurately and efficiently reviewed by an operator if required. The proposed approach allows a reliable detection of waterborne micro-organisms in large quantities of water. This provides an industrial system to monitor micro-organisms in the water industry. The implemented algorithm complies with the standard operating protocol provided by the water authorities in UK and satisfies the EU directive on drinking water quality

The MV second example presented in this chapter can improve the performance of automated analysis to determine whether suspicious activity is present in an open scene under CCTV surveillance. The system can determine the presence and movements of a person and provide indications of suspicious activity based on pathway taken, speed, direction; the system can also provide indication of suspicious activity based on human posture analysis. These enable efficient monitoring and accurate review of scenes by an operator. The system proposed can provide a warning role to reduce the problem of human operator's fatigue and shortened attention span. This greatly increases the ability to carry out properly the task of constant and distant monitoring.

9. Acknowledgment

The author wishes to thank the Scottish Parasitic Laboratory for providing samples of micro-organisms in water. Also I would like to thank my wife Vicky Grandon for her continual support and comprehension. Special thanks to my students Sorin Bota and John Freer.

10. References

BIDS (2011) Bath Information and Data Services, available from
        http:// www.bids.ac.uk

Braggins, D. (2000). Illumination for machine vision. *Sensor Review*, Volume 20, No1, pp 20-30.

Butler, B. E. (1980). Selective Attention and Stimulus Localization in Visual Perception. *Canadian Journal of Psychology*, Vol. 34, 1980 pp 119-133

Current, W.L. (1988). The biology of Cryptosporidium. *Am. Soc. Microbiology News*, 54 / 1988, pp. 605-611

Current, W.L. And Gracia. (1991). Cryptosporidiosis. *Clinical Microbiology Review*, 4/1991, pp. 325-358

Danielsson, P. (1980). Euclidean distance mapping. *Computer Graphics and Image Processing*, 14/1980, pp.227-248

EEC (1998). Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption, *Official Journal* L 330 , 05/12/1998 pp. 0032 - 005

Ellenberger, S. & Ian T. Young, I. (2000). Microscope image acquisition, Image Processing and Analysis, Oxford University Press, pp.1-35

Eriksen, C. W., Murphy, T. D. (1987). Movement of attentional focus across the visual field: A critical look at the evidence., *Perception & Psychophysics*, Vol. 42, 1987 pp 229-305.

Eriksen, C. W., Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays, *Perception & Psychophysics*, Vol. 12, 1972 pp 201-204.

Fernandez- Canque, H. L.; Bota, S.; Beggs, B.; Smith, E.; Hintea, S. and Smith, H. V. (2000). Detection and Classification System for Water-Borne Micro-organisms Using Machine Vision. *Journal TABEMED 2000 Acta electrotehnica napocensis.* Romania. June 2000 ISSN 1224-2497 pp 105-111

Fernandez-Canque, H. L.; Hintea, S.; Csipkes, G.; Bota, S. and Smith, H. (2009). Machine Vision Application to the Detection of Water-Borne Micro-Organisms. *Intelligent Decision Technologies an International Journal.* IOS Press, ISSN 18724981,2009 Vol 3, Number 2. pp 93-100.

Fernandez-Canque, H. L.; Hintea, S.; Freer, J. And Ahmadinia, A. (2009). Machine Vision Application to Automatic Intruder Detection using CCTV. *Proceeding 13th International Conference, KES2009 Knowledge-Based Intelligent Information and Engineering Systems.* ISBN 13978364204594, Santiago, Chile September 2009. pp. 498-505.

Frazen, C and Muller, A. (1999). Cryptosporidia and Microsporidia – Waterborne Diseases in the Immunocompromised Host. *Diagn Microbiol Infect Dis*, 1999, 34: 245-262

Hanks, J. (1998). Using Edge Detection in Machine Vision Gauging Application. , *National Instruments* - 1998. Application Note 125.

Goodman, J.W. (1996). Introduction to Fourier optics, 2nd edition, McGraw-Hill, San Francisco, USA.

Goolkasian, P., (1991) Processing Visual-stimuli Inside and Outside the Focus of Attention. *Bulletin of the Psychonomic Society*. 1991 Vol. 29, No. 6, p 510.

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* 82 (1): pp 35–45.

Luo, X., Jayas, D.S. and Symons, S. J.(1999) Identification of Damaged Kernels in Wheat using a Colour Machine Vision System, *Journal of Cereal Science.* 1999. 30 pp 49-59.

Pawley, J.B. (1995). Handbook of biological confocal microscopy, 2nd edition, Plenum Press, New York.

Russell, S. J.; Norvig, P. (2003), Artificial Intelligence: A Modern Approach 2nd ed., Upper Saddle River, New Jersey: Prentice Hall, pp. 111–114, ISBN 0-13-790395-2

Saarinen, J., Julesz, B. (1991) The Speed of Attentional Shifts in the Visual field. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 88, No. 5, pp 1812-1814.

Smith, H.V. and Rose, LB. (1990). Waterborne Cryptosporidiosis. *Parasitol Today* 1990. 6, pp. 8-12.

Smith, H.V. and Rose, LB. (1998). Waterborne cryptosporidiosis: Current status. *Parasitol Today. 1998.* 14, pp.14-22.
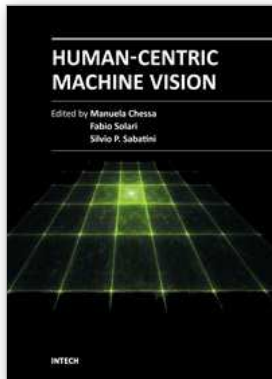
Human-Centric Machine Vision

Edited by Dr. Fabio Solari

ISBN 978-953-51-0563-3

Hard cover, 180 pages

Publisher InTech

Published online 02, May, 2012

Published in print edition May, 2012

Recently, the algorithms for the processing of the visual information have greatly evolved, providing efficient and effective solutions to cope with the variability and the complexity of real-world environments. These achievements yield to the development of Machine Vision systems that overcome the typical industrial applications, where the environments are controlled and the tasks are very specific, towards the use of innovative solutions to face with everyday needs of people. The Human-Centric Machine Vision can help to solve the problems raised by the needs of our society, e.g. security and safety, health care, medical imaging, and human machine interface. In such applications it is necessary to handle changing, unpredictable and complex situations, and to take care of the presence of humans.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Hernando Fernandez-Canque (2012). Machine Vision Application to Automatic Detection of Living Cells/Objects, Human-Centric Machine Vision, Dr. Fabio Solari (Ed.), ISBN: 978-953-51-0563-3, InTech, Available from: http://www.intechopen.com/books/human-centric-machine-vision/machine-vision-application-to-automatic-detection-of-living-cells-objects

INTECH
open science | open minds