

# A Review of Computer Vision Techniques in the Detection of Metal Failures

Deborah Fitzgerald<sup>1\*</sup> and Roselita Fragoudakis<sup>1</sup>

<sup>1</sup>Department of Mechanical Engineering, Merrimack College, North Andover, Massachusetts, USA

**Abstract.** This paper considers and contrasts several computer vision techniques used to detect defects in metallic components during manufacturing or in service. Methodologies include statistical analysis, weighted entropy modification, Fourier transformations, neural networks, and deep learning. Such systems are used by manufacturers to perform non-destructive testing and inspection of components at high speeds [1]; providing better error detection than traditional human visual inspection, and lower costs [2]. This is a review of the computer vision system comparing different mathematical analysis in order to illustrate the strengths and weaknesses relative to the nature of the defect. It includes exemplar that histograms and statistical analysis operate best with significant contrast between the defect and background, that co-occurrence matrix and Gabor filtering are computationally expensive, that structural analysis is useful when there are repeated patterns, that Fourier transforms, applied to spatial data, need windowing to capture localized issues, and that neural networks can be utilized after training.

## 1. Introduction

Computer Vision encompasses several technologies, including image processing and algorithm development to make decisions or interpretations based on the image data. Computer Vision is used in numerous applications including robotics, autonomous vehicles, medical imaging, and manufacturing. Such systems are used by manufacturers to perform non-destructive testing and inspection of components at high speeds [1]; they provide better error detection rates than traditional human visual inspection, can lower costs [2], and may operate in a broader range of the electromagnetic spectrum than human vision [3]. Analysts predict by 2027, the global computer vision market will be worth 19 billion USD [4]. Computer Vision, also called Machine Vision in industrial applications, continues to evolve and play a major role in manufacturing inspection, unit identification, measurements, and guidance for automated part placement [5].

The study of Computer Vision began at MIT in the 1960s with Larry Roberts' thesis discussing the possibilities of extracting 3D geometrical information from 2D perspective view of blocks (polyhedral) [6]. Another MIT pioneer was David Marr, who in 1978 proposed a framework to understand real-world scenes by moving from low-level edge detection of 2D images to high-level modeling of 3D objects [6]. Since Marr, researchers have developed statistical analysis, structural analysis, filtering methods transforms, neural networks, and deep learning [2]. As computing power has grown, Computer Vision applications have been utilized outside of research facilities.

### 1.1. Overview of Computer Vision System and Digitization

A Computer Vision system consists of physical components, such as a light source, a lens, and a camera, as well as vision software which performs image manipulation, pre-

---

\*Corresponding author: [fitzgeraldde@merrimack.edu](mailto:fitzgeraldde@merrimack.edu)

processing, segmentation, classification, and interpretation [7]. The simple image collection system shown in Figure 2, uses a light source to reflect off an object and a sensor to measure the intensity of the reflected light [8]. The software requires images to be digitized.

An analog image is a 2D image referred to as  $F(x,y)$ . The function  $F$  is used to denote the infinite precision of the image in spatial parameters,  $x$  and  $y$ , as well as its infinite precision in intensity at each spatial point  $(x,y)$ . On the other hand, a digital image is a 2D image referred to as  $I[r,c]$ , and is represented by a discrete 2D array of intensity samples. Each sample is represented using a limited precision [8]. Picture elements, or pixels, are obtained by sampling the continuous image at different  $(x,y)$  points. The sampling frequency and number of samples will determine the accuracy.

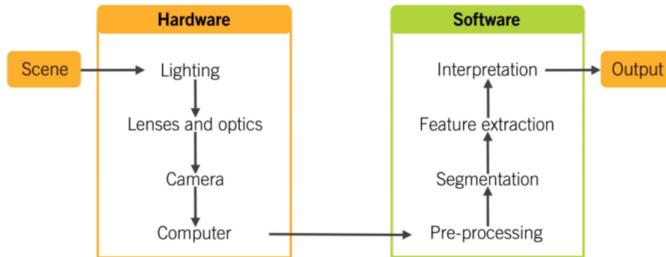


Fig. 1. Elements of a Computer Vision System [7]

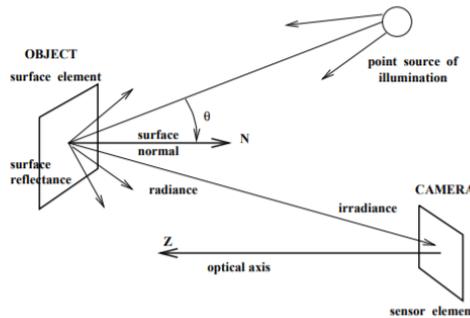


Fig. 2. Simple Image Collection [8]

There are two well-known representations for images: grayscale and multispectral [8]. The first stores one intensity value for each pixel; grayscale uses black to represent zero reflected light, white to represent the total reflection of light, and variations of gray to capture all other intensities. The second stores a vector for each pixel [8]; for color images, this vector can be based on brightness and the three values per pixel map to red, green, blue (RGB) or on reflected light and map to cyan, magenta, and yellow (CMY).

Computer data storage is defined in terms of the number of bits or bytes required; eight bits can represent integer values ranging from 0 to 255. Thus, image resolution can be defined in terms of the number of pixel rows (height) and columns (width) and the data size for each pixel can be used to determine the memory requirement. A matrix can represent the different grayscale value at each position and would take the form of equation 1 [9].

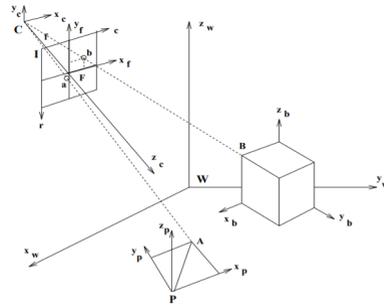
$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,M-1) \\ f(1,0) & f(1,1) & \dots & f(1,M-1) \\ f(2,0) & \vdots & \vdots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,M-1) \end{bmatrix} \quad (1)$$

## 1.2.Edge Detection

Edges of physical objects can become blurred once an image has been digitized. Therefore, researchers have developed logic to try to determine when a transition occurs [7]. This is important when analyzing an image with multiple objects and when differentiating an object from its background.

## 1.3.Coordinate Systems

The Computer Vision analysis may not require the full rendering of the 3D scene as proposed by Marr; in some cases, the decision can be made using only edge detection or a subset of the full model [6]. There are several coordinate systems that correspond to an image and scene: “world W, object O (pyramid Op or block Ob), camera C, real image F and pixel image I” [8]. Figure 3 shows the different coordinate systems for two objects P and B [8]. When computer vision is used for inspection, there may be a relationship between one or more objects (e.g. engine block and missing bolt) or the scene may only have a component and its background.



**Fig. 3.** Coordinate Systems used in modeling and analysis [8]

## 2.Methodologies and Algorithms Used in Computer Vision

There are different algorithms and ways to analyze and interpret the digital data. Truong mentions several successful visual inspection methods. Some algorithms are better for a specific material (e.g. paper, steel) [1] or industry. Others, like structural analysis, work well for materials with a repeating structure or pattern [2].

The methods that are discussed include: statistical analysis using Otsu’s method and Truong’s weighted entropy modification, Fourier transformations, and Song’s neural networks; these methods can be applied for inspection and defect detection.

### 2.1.Statistical Analysis

Statistics and probability can be used to generate histograms and calculate divergence, normalizations, and chi-square. Histograms can be unimodal, bimodal, and multimodal.

#### 2.1.1.Otsu’s Method

In 1979, Otsu defined a method of automatic thresholding; in which a value is chosen such that the image can be divided into background and objects; it converts grayscale images into k-ary images [1].

Consider the Image Matrix of equation 1, in which the value  $f(x,y)$  is the grayscale value at the position  $(x,y)$ . The system has  $L$  gray values; thus, the pixel range is from 0 to value  $L-1$  [1]. Equation 2 specifies the histogram, while equation 3 gives the average. The

threshold value  $k$  is then selected and used to generate equations 4 and 5, which are used to specify the foreground  $C_1$  and background  $C_2$  [1]. Equations 6 and 7 give the probability of foreground and its corresponding average. Finally, equations 8 and 9 give the probability of the background and its corresponding average [1].

Let the number of pixels with gray value  $i$  be  $n_i$  and the total number of pixels in  $I$  be  $N$

$$p_i = \frac{n_i}{N} \quad (2)$$

Then, the average gray value of  $I$  is calculated by:

$$u_1 = \sum_{i=0}^{L-1} i p_i \quad (3)$$

Due to the thresholding, the pixels of  $I$  are divided into two classes

$$C_1 = \left\{ (x, y) \mid \leq f(x, y) \leq k \right\} \quad (4)$$

$$C_2 = \left\{ (x, y) \mid k + 1 \leq f(x, y) \leq L - 1 \right\}$$

The probability of class occurrence and the average gray value of each class are based on

$$w_1(k) = \sum_{i=0}^k p_i \quad (5)$$

$$w_2(k) = \sum_{i=k+1}^{L-1} p_i \quad (6)$$

$$u_1(k) = \sum_{i=0}^k \frac{i p_i}{w_1(k)} \quad (7)$$

$$u_2(k) = \sum_{i=k+1}^k (L - 1) \frac{i p_i}{w_2(k)} \quad (8)$$

The optimal threshold,  $k^*$ , must maximize between class variance

$$k^* = \arg \max \left\{ w_1(k) (u_1(k) - u_n^2) + w_2(k) (u_2(k) - u_n^2) \right\} \text{ for } 0 \leq k + 1 \leq L - 1 \quad (9)$$

Otsu's method successfully classifies background and foreground for images with certain characteristics, namely the probability histogram is binary or multimodal [1]. Figure 4 shows the defect (a), the classification (b), and (c) the histogram [1]. The result does not detect the defect. According to Truong, histograms from real-life images have multimodal distributions as the objects are usually larger compared to image size [1]. However, when

images are obtained from visual inspection systems, their histograms tend to have unimodal distributions due to the uniformity in luminance and size of the defect regions, thus preventing Otsu's method to identify the desired threshold values [1].

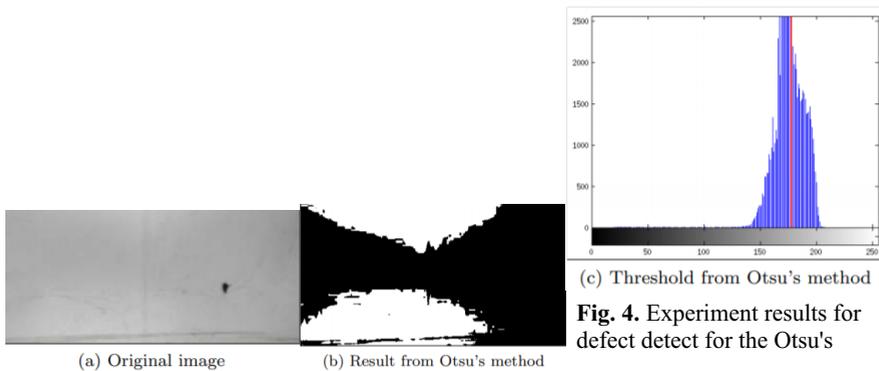
### 2.1.2. Weighted Otsu Method

Several researchers tried to fix Otsu's method by adding a weight to the optimization equation, as for example the use of a complement of the probability occurrence  $(1-p_k)$  [1].

$$k^* = \arg \max \{ \psi(k)(w_1(k)U_j^2(k) + w_2(k)(U_2^2(k))) \} \text{ for } 0 \leq k + 1 \leq L - 1 \quad (10)$$

Truong tried to calculate an entropy function of the image and then used this as the weight multiplier. Entropy measures the *statistical randomness* of the image and therefore, he reasoned that in images without surface defects there would be small levels of entropy and those with defects there would be large values. Thus, Truong used the entropy work by Kapur and the classification work by Otsu to generate a new algorithm; the  $\psi$  is the entropy function [1].

$$k^* = \arg \max \{ \psi(k)(w_1(k)U_j^2(k) + w_2(k)(U_2^2(k))) \} \text{ for } 0 \leq k + 1 \leq L - 1 \quad (11)$$



**Fig. 4.** Experiment results for defect detect for the Otsu's

### 2.1.3. Co-occurrence Matrices

Another statistical method, called the co-occurrence matrix, uses the work done by Haralick in 1973 and Connors and Harlow circa 1980. The technique considers pairs of pixels in certain spatial relations to each other [10]. The technique calculates the probability of the grayscale intensity of two pixels to be the specified values  $(i,j)$  when they are at a certain distance, as measured in polar coordinates  $(d, \Theta)$ ; thus, it effectively considers the relationship of grayscale of neighborhood of pixels as shown in figure 5. Further, it calculates the energy, entropy, correlation, local homogeneity, and inertia of the results to classify the texture, which allow to analyze the image texture features, using also the co-occurrence matrix.

In figure 5 the distance  $D$  and the value of  $\Theta$  determine the diagonals from pixel  $(n_x, m_x)$  [11]. Point  $(D_x, D_y)$  uses the following polar relations

$$D_x = D \cos \Theta \quad (12)$$

$$D_y = D \sin \Theta \quad (13)$$

The co-occurrence matrix, CM, for the image  $I(k,k)$  is defined by [11]

$$C_M = \sum_{n=1}^k \sum_{m=1}^k \begin{cases} 1 & \text{if } I(n, m) = k \text{ and } I(n + D_x, m + D_y) = k \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

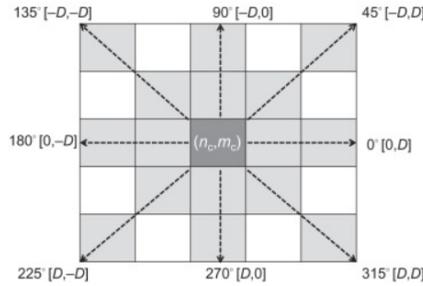


Fig.5. Determination of Co-occurrence Matrix

## 2.2. Structural Analysis

Certain materials exhibit patterns. If there is a defect, then the pattern will not repeat. Thus, structural analysis can be used to detect defects for materials or products that have well known patterns. [2]

## 2.3. Fourier Transform and Gabor Filtering

To decompose an image into its trigonometric components ( $\sin \Theta$  and  $\cos \Theta$ ) Fourier transforms are used. Gabor Filtering is a windowing mechanism used in conjunction with Fourier Transforms. The Fourier Transform focuses on broad area and the Gabor can narrow the scope of the area under consideration; it localizes the analysis [2]. For a square image of size  $N \times N$ , the two-dimensional DFT is given by the Fourier transform in Equation 15; the inverse Fourier transform is shown in Equation 16.

$$F(k, l) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) e^{-i2\pi \left( \frac{k_i}{N} + \frac{l_j}{N} \right)} \quad (15)$$

$$f(a, b) = 1/N^2 \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} F(k, l) e^{i2\pi \left( \frac{k_a}{N} + \frac{l_b}{N} \right)} \quad (16)$$

## 2.4. Neural Networks and Deep Learning

Computer Science has developed some concepts like neural networks and deep learning. These algorithms and mechanism that can be applied to problems in Computer Vision. A neural network consists of one or more hidden layers, used for the understanding go complex structures of data [12]. The algorithms created look for patterns in such data sets, rather than having a set of specific rules when analyzing the input data. This paradigm

requires training of the system with example patterns. Machine Learning systems and Neural Networks are trained using supervised, unsupervised, or semi-supervised learning.

#### *2.4.1. Convolution Neural Networks, Fully Convolutional Networks, U-net, DU-net*

In Convolution Neural Networks, the input image is convoluted with a smaller matrix and the resultant is used to look for patterns. Deka explained how CNN was used to detect errors. The system was trained using a set of examples of good and bad images [12].

Additional research has been done to improve the CNN detection. According to Song, while the traditional convolution neural network classifies images, it is unable to realize segmentation [2]. The Fully Convolution Networks (FCN) added a step in which the background is segmented from the target area; this leads to better results because the system focuses the analyzation on the target area [2].

Song created a CNN that focused on pixel analysis, rather than image analysis; it builds upon Long's FCN and U-net. The DU-net system, the result of the previous step is fed into the next step; consisting of either applying a convolution, deconvolution or a pool merge operation. The U-net and DU-net require less training data than FCN [2].

### **3. Computer Vision Techniques Used in Metal Failure Detection**

Understanding the deformation response in the near crack-region of metallic specimens undergoing fatigue loading may give valuable information on the failure process as fatigue cracks propagate, as well as help in the identification of phenomena such as crack closure and partial crack opening. Gonzales et al., use 3D-DIC analysis to assess the aforementioned cyclic plastic deformation response, and related phenomena [13]. The use of DIC allows for an assessment of the displacement field in very small length scales and consequently the calculation of the associated strain and potentially stress fields in the region of concern [14]. A combination of stereo-DIC and Scanning Electron Microscopy (SEM) techniques, have been used successfully by Tong et. al., in reporting near-tip strain ratcheting under cyclic loading [13].

Additive manufacturing may be an established manufacturing process, however, it is still one that is under constant development, especially in the case of metallic raw materials. Similar to all manufacturing processes, quality control is one of the most crucial steps in this process, and may be the step where failure detection is the key component of the final quality of the product. Zhu et al. discuss a combination of X-ray computed tomography, computer vision, and machine learning as an inspection pipeline in detecting and understanding pore evolution in post-processing of binder jetting materials [15].

Chemical corrosion on metallic surfaces is a dangerous failure mechanism and a complex problem to identify and analyze. Enikeev et al. suggest a step by step algorithm for processing chemical corrosion data that combines image processing, image binarization, and identification of object contours and analysis of object characterization. Their algorithm is based on fractal analysis for corrosion of cracking specimens, giving thus, a more complete analysis and representation of the failure the metallic surface is subjected to [16-17].

### **4. Conclusion**

Computer Vision can be used by manufacturers to detect errors, ensure tolerances, and perform inspections. Deployed systems detect scratches, indentation, and strains in metals. Methodologies under study combine DIC, X-ray, and Computer Vision for enhanced detection. Computer Vision employs different algorithms: statistical analysis, structural analysis, transforms, and Neural Network/Deep Learning. Statistical analysis works well in

scenarios in which there is high contrast between the defect and background. Structural analysis focuses on repeated patterns. Fourier transforms are applied to spatial data and combined with windowing techniques. Neural Networks on the other hand, require training. Computer Vision is an evolving technology; research in Machine Learning and Artificial Intelligence will continue to improve the algorithm capabilities and the detection of faults in metals.

## References

1. M.T.N. Truong, S. Kim, Automatic image thresholding using Otsu's method and entropy weighting scheme for surface defect detection, *Soft Comput* **22**, 4197-4203 (2018)
2. L. Song, W. Lin, Y. Tang, X. Zhu, Q. Guo, J. Xi, *IEEE Access*, **7**, 27547-27554 (2019)
3. B. Dipert, *Beyond-visible Light Applications in Computer Vision, Embedded Vision Alliance* (2017)
4. A. Ajgaonkar, K. Seier, *The Value of Computer Vision: more Than Meets The Eye, Insight Tech Journal, Spring* (2021)
5. COGNEX Corp., Introduction To Machine Vision: A guide to automating process & quality improvements, 6 (2016)
6. T.S. Huang, Computer Vision Evolution and Promise, CERN-96-08, 21-27, (1996)
7. B. Anthony, MIT Professional Education: Smart Manufacturing. Week 6 Slides, (2019)
8. L.G. Shapiro, G.C. Stockman, *Computer Vision*, 33-60, (2001)
9. A.K. Maini, *Image Processing Using MATLAB: Basic Operations*, (Part 1 of 4) electronicsforu.com (2019)
10. E.R. Davies, *Computer Vision: Principles, Algorithms, Applications, Learning*, 190-192 (2018)
11. L.S. Athanasiou, D.I. Fotiadis, L.K. Michalis, 4- Plaque Characterization Methods Using Intravascular Ultrasound Imaging, *Atherosclerotic Plaque Characterization Methods Based on Coronary Imaging*, Academic Press, 71-94 (2017)
12. P. Deka, R. Mittal, Quality inspection in manufacturing using deep learning based computer vision: Improving yield by removing bad quality material with image recognition, *Toward Data Science*, December (2018)
13. G. Gonzáles, J. González, et al. Measuring elastoplastic strain loops in the near crack-tip region using a Stereo Microscope DIC system, *Int J. Fatigue*, **133** 105427 (2020)
14. M.A. Sutton, J.J. Orteu, H. Schreier, *Image Correlation for Shape, Motion and Deformation Measurements: Basic Concepts, Theory and Applications*. Springer (2009)
15. J. Tong, B. Lin, et al. Near-tip strain evolution under cyclic loading: In situ experimental observation and numerical modelling. *Int J. Fatigue*, **71**, 45–52 (2015)
16. Y. Zhu, Z. Wu, W.D. Hartley, et al. Unraveling pore evolution in post-processing of binder jetting materials: X-ray computed tomography, computer vision, and machine learning, *Additive Manufacturing*, **34**, 101183, (2020)
17. M. Enikeev, I. Gubaydullin, M. Maleeva, Analysis of Corrosion Process Development on Metals by Means of Computer Vision, *Engineering Journal* **21**, 183-192 (2017)