Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 38 (2019) 408–414

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

29th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2019), June 24-28, 2019, Limerick, Ireland.

Object recognition within smart manufacturing

Andrew D O' Riordan[a,b], Daniel Toal[a,b], Thomas Newe[a,b], Gerard Dooly[a,b]

[a]Centre of Robotics & Intelligent Systems, University of Limerick, Ireland
[b]Confirm Research Centre, University of Limerick, Ireland

Abstract

Process automation has become a norm within industry with cheap and easily accessible automation technology becoming a standard option available to manufacturing firms. However, without the implementation of flexibility in manufacturing by the employment of intelligent systems, the technology will be limited in application. The development of smart sensing technologies has allowed for modularity and versatility to become familiar terms on a manufacturing floor, notwithstanding, it is still widely recognised that a human employee is the most valuable and flexible asset a company may have. Automation falls short in terms of flexibility due to its lack of independence during operations with high levels of variance, such as varying target position from cycle to cycle. Processes with high levels of variance disallow employment at a satisfactory level of standard or more traditional automation methods due to the lack of ability of current systems to deal with the unexpected. This paper aims to examine the technology used, and that can be potentially used in processes with high levels of variance, specifically, vision systems used in collaboration with an algorithmic comparison to compare an obtained image to an image or 3D model of the target for target recognition/object identification.

While there have been copious experiments to employ imaging technology for object identification, some systems do currently occupy the factory floors of manufacturing facilities for recognition. Many of these systems run on RFID, barcode or fiducial marker. These technologies, while operational, require a pre-emptive effort to be made to ensure all products or objects to be recognised have an identifying marker attached before recognition is possible. This substantially limits the flexibility of manufacturing as the versatility of a process line to adapt to different products needs to be a possibility without the necessity of rebranding or retagging each object or product, causing a decelerative rate of production.

The first section in this paper identifies the most commonly used methods of object recognition and the necessary modules required for each different algorithmic architecture. The need for particular architectures, depending on object illumination, shape, texture etc., will be addressed in the second section, as well as the differences in the construction of these architectures by use of different combinations of modules. Finally, the third section aims to address the best industrial practice and the opportunities currently being offered by research. It is the aim of the authors to determine where there are gaps to be filled between industry and research. This paper will identify areas of research that need to be examined in order to close the gap between theory and practice in research and industry respectively, to allow industry 4.0 to become a reality across the board for manufacturing

Peer-review under responsibility of the scientific committee of the Flexible Automation and Intelligent Manufacturing 2019 (FAIM 2019)
10.1016/j.promfg.2020.01.052

Peer-review under responsibility of the scientific committee of the Flexible Automation and Intelligent Manufacturing 2019 (FAIM 2019)

Keywords: Object Recognition; Machine Vision; Target Recognition; Smart Manufacturing

1. Introduction

Object recognition (OR) is essentially a challenge of matching the shape, illumination, reflectivity/texture and orientation of an acquired image to an existing image. This technology is currently implemented in industry on a basic level of visual inspection, where an image will be captured of an object and compared to an ideal image stored as a reference. Although well-established within manufacturing of metallurgic parts for surface inspection for corrosion, this method has shown to be successful in the agricultural industry allowing for quality assurance of both wheat and apples [1][2]. Another clear demonstration of the capabilities of modern vision systems for visual inspection is the ability of modern stereo vision systems in image matching techniques. Many modern image matching techniques in stereo vision utilise in-depth pixel analysis, shrinking window analysis and coarse-fine hierarchies [3]. The advancement of technology associated with or designed for operation on visual images (both within and outside of the visible spectrum) has attributed to the ability of OR as a research area to grow. The challenge originally faced by OR technology was to grasp the edge of an object to differentiate it from the background and to identify which object it was from a set of predetermined options. While many researchers in the earlier stages of object detection overcame the first half of the challenge (i.e. finding the object within the image) with what are considered standard algorithmic solutions today (such as background erasure by use of segmentation and thresholds [4])]). More modern research has found that the challenge is successfully addressed through feature extraction and subsequent object classification. Within the feature extraction portion of the procedure, it is necessary for the image acquired to pass through a number of modules within the architecture of the algorithm. Within this architecture, the modules differ from procedure to procedure as due to the nature of the images acquired or the object to be recognised; there exists a necessity for particular filters, nonlinear operators and pooling operations to take place. Generally, the first module to operate is a filter bank. In some operations the filter bank may be repeated in a second stage of a feature extraction operator, using edge detectors or a group of subsequent filters using gradient descent [5].

Following the capture of the image and the identification of the object as an independent body within the image, an operation has to occur which classifies the object through a comparison to a pre-existing dataset based on the characteristics of the object itself. From the filter bank, the output will be passed to a non-linear operation to normalise the output. Finally, the pooling operation takes place This can be carried out in a number of ways including a comparison of the object as identified from the target image, to the object as identified within images as part of a dataset (usually numerous poses and illuminations are contained within this dataset) or is often based on nearby values of real space or feature space [5].

2. Modules of Algorithm Architecture

OR can best be described as being based on a two-stage process; Feature Extraction and Classification. While single stage systems do exist, more systems are employing two or more successive stages. As addressed in section 1, there are multiple modules that need to be pieced together for OR to execute successfully; these modules combine to make these two stages. Although there is no strict set order for the modules to occur, there is a consensus in research that they run in a specific order (as addressed in 3). This order allows the output of one set of modules to be optimised by the subsequent set, unless where it has been found beneficial to recycle a module output through a module a second time. This section will identify the modules, their make-up and purpose within the subsections of

2.1. Feature Extraction

Feature extraction, being the most important and fundamental research question about object and pattern recognition, has had much research focused on it. The primary objective is to extract important features from image data in order to create a comparison between the data obtained and a predefined dataset. Many earlier feature extraction techniques utilised information from the image such as visual features, edge detection, pixel analysis or algebraic features. These focused mostly on boundaries of shapes within the image, disregarding features such as patterns within the shape. In more recent research, however, the patterns within a shape have proven to be integral to the advancement of OR, these patterns include facial features, hatching detail etc. [6].

The purpose of a filter bank is to discriminate between points of interest within an image and those not of interest. Filter banks operate on the principle of separation of the input signal from the acquired image into multiple components. These components are subject to a set of thresholds (sometimes predetermined and sometimes calculated on the image captured as a whole based on statistical analysis) to eliminate anything that does not stand out from the background of the image as a feature. Though commonly operated based on oriented edge detectors, such as SIFT or HOG, filter banks often utilise Gabor Wavelets [6] [7]. The difference in the execution of various filter banks is due to the discontinuities that exist, particularly during observation of a signal at a small window length [8]. Filter banks are composed of convolution filters, gain coefficients and a sigmoid/tanh non-linearity. Filter banks give an output which can then be used for identification of features, however, due to the nature of the filter bank output, a non-linear operation must be carried out.

The non-linear operation often involves a rectifier and another operation. The rectifier applies an absolute value function to its input. The normalising layer, widely used as the second layer within a non-linear operation, will account for variance of contrast or brightness within images for comparison purposes [5].

The pooling layer is used to resolve issues due to small distortions. This is achieved by pooling together filters that have extracted features with high similarity, often carried out over max or averaging operations [9]. Pooling operations are generally carried out on adjacent windows within a neighbourhood, where the window size is a predetermined n X n group of pixels [10], however, have also proven successful when applied to dimensional spacing based on feature types or spatial dimensions [5].

2.2. Classifiers

Although generally associated with machine learning, classifiers are employed within OR to determine what outputs from the feature extraction are to be utilised. This can be carried out in unsupervised learning or supervised learning fashion [11] [10]. Learning, in relation to classification operations determine the relationship between the inputs and outputs of feature recognition in order to streamline the process moving forward and re-employ the most successful path from input to output [12]. Below, both supervised learning and unsupervised learning are discussed in relation to classifiers.

2.3. Supervised Learning

Supervised learning occurs when the input and output of the feature extraction operations are known. Where the relationship between an input and an output is f(x)=Y, the function of supervised learning is where x (input) and Y (output) are known, and f(x) must be determined by the classification operation. This determination must be based on a dataset obtained (x1,y1...x2,y2...xn,yn...etc.) [13]. The challenge encountered when using supervised learning is the fact that all data must be labelled, the algorithms will not inherit structure based on input alone. Supervised learning, therefore, will only determine the characteristics of an existing function [14].

2.4. Unsupervised Learning

Unsupervised learning differs from supervised in that the input data is known. However, there is a lack of corresponding output data. A prominent aim for unsupervised learning is to determine a model for the preceding structure or data distribution and to identify trends etc. to learn about the data. Furthermore, algorithms are left to identify the trends in data and to employ this information [12]. Unsupervised learning within OR works to identify and learn the filter and pooling operations that cluster multiple filter outputs together [10]. The classifications from unsupervised learning can be employed again through the use of the relationships found, with the application of these relationships to patches of the image again, to identify more invariant features [15].

3. Architecture Types

The Stereotypical Architecture for OR can be broken down as follows; An image is acquired, this image is then passed through filter banks. The output of these filter banks is then subjected to a non-linear operation and subsequently to a normalisation process. The outputs, having been normalised, enter a pooling operation. The output of the pooling operation is classified by either supervised or unsupervised learning classifiers. The output of the classifier can then be applied to the image again for accurate recognition to occur, or, can be applied to a known dataset for determination of object identification, pose and distance [5].

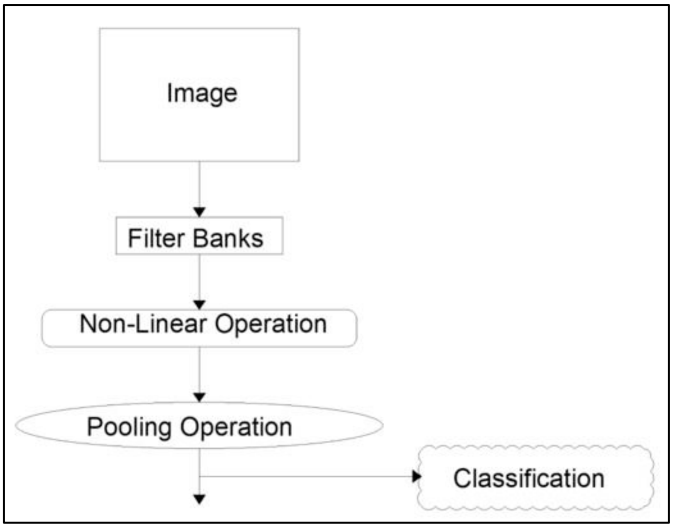

Figure 1 Visual Algorithm Architecture

Many variations of this operation exist, however, and have proven successful when tested against Caltech-101 (a benchmark dataset). Many authors have employed two stages of learned feature extractors; both made up of filter banks, non-linearities and pooling. This involved both supervised and unsupervised learning for convolutional networks. This yielded results from 50 to 60 percent accuracy [5] [15] [16].

4. Object Recognition within Industry 4.0

Industry 4.0 is a major target for both industry and academia. Described as the 4th industrial revolution, it will be the first industrial revolution that has been pre-planned. The first industrial revolution was aided with the utilisation of steam power, the second with electricity and the third with automation, Industry 4.0 is based on the application of digitisation and smart manufacturing. This step will be driven by the connection of physical world elements with the industrial internet of things (IIOT). The capture of data, analysis and application of results within this data, (from products, processes and production systems) will aid the step toward Industry 4.0 majorly [17] [18]. This section aims to address the opportunities and impacts of OR within Industry 4.0 and the challenges of implementation of it on a large scale.

To address the opportunities that the employment of OR may present, first the objectives of industry 4.0 must be examined. Identified as the major pillars of industry 4.0 are;

- Autonomous robotics
- Simulation
- Horizontal and Vertical System Integration
- IIOT

- Cybersecurity
- The Cloud
- Additive Manufacturing
- Augmented reality
- Big Data Analytics [19].

Within these pillars, this paper will identify opportunities and impacts caused by OR in the following four;

- Autonomous robotics
- Cloud Computing Services
- IIOT
- Big data analytics.

4.1. Autonomous Robotics

Autonomy within robots is a hugely important factor in Industry 4.0. The current industry standard is robotics completing a basic repetitive task; however, it is based on location and therefore does not allow for variation. Where a robot may need to pick up a part for a process to begin, the part must be in a predetermined set of cartesian coordinates relative to the robot. This disallows flexibility and autonomy within pick and place operations or feed operations. This predicament has led to humans having to be employed in many facilities to carry out mundane tasks which are well within the scope of capability of robotics but are unavailable to standard robotics due to a level of variance in part placement. The employment of OR would allow for robotic operations to take place based on the information gathered on the object pose and distance from the camera, meaning that human employees may be freed up in order for their potential to be applied to more complex tasks. This would also allow for MES tracking of each part as it is sensing the object to be within certain workstations etc. [20].

4.2. Cloud Computing Services

Due to the high volumes of data to be communicated, large, powerful servers will need to be dedicated to OR. However, due to the fast-paced nature of manufacturing, the necessity for a company to own its own servers is not consistent. The opportunity for a company to use remote servers not only removes the capital investment risk, but it also allows floor space and resources that would have been dedicated to large servers to be used as a part of the production environment. Therefore, the decentralisation of decision making and analysis from the local device to the cloud is imperative moving forward (i.e. OR algorithms may not run locally but rather in the cloud). Data centres dedicated to housing cloud computing facilities is expected to rise steadily. Cloud services will need to be secure against both physical and cyber threats in order to ensure that the flow of data to and from the cloud is not interrupted resulting in downtime on factory floors [21] [22].

4.3. IIOT

IIOt is the Industrial Internet of Things, defined as the collection of all industrial related equipment connected via internet. IIOT is expected to grow to over 50 billion devices interconnected by 2020 [23]. The infrastructure necessary to carry the inputs and outputs of each of these devices intended to be communicated over the IIOT can only be accommodated by investment and development. This is particularly true in relation to OR. With OR being employed, the dataset for image comparison will be too large to store locally at the camera. This creates the necessity for storage of this dataset on a remote server (i.e. cloud). The constant streaming of images from each camera within a facility carrying out OR generates a demand for reliable internet services, with low latency and large longevity. Bandwidth must also be accommodating to the communication of this data [24]. IIOT will, therefore, have to be ever-changing and dynamically improving to accommodate OR as a technology.

4.4. Big Data

A major benefit of digitisation within industry is the opportunity to employ machine learning in order to optimise future procedures. This, however, requires a dataset to be observed. This dataset in relation to OR will consist of multiple images of multiple angles and illuminations for each object to be recognised, resulting in potentially 10's of thousands of images being stored. There will also be a need to store the history of procedures and their success rates (i.e. unsupervised and supervised learning datasets). The requirement for this storage can be met via off-site data centres, as discussed in 4.2. Similarly, there is also the need for constant expansion as more products are digitised there is more and more heterogeneous data to be collected. OR, if employed effectively within industry, would create a huge requirement for more dedicated storage within data centres. The employment of analytics to datasets as large as this would allow a significant decrease in creative design times on products or procedures and an increase in cycle efficiency comparisons [25] [26].

5. Conclusion

There is great potential for the employment of OR; however, due to the dense scenes within factory floors, more research is needed into the employment of filters that can discriminate between objects of interest and background features. Another problem to be addressed is the handling of the lack of consistency of lighting within manufacturing environments as within a dynamic environment, light and shading can frequently differ due to general operations. OR being rolled out within industry would have a huge impact on product cycle times, cost and reliability, however, this is not without cost. Due to the investment needed for infrastructure upgrades and upkeep within the factory, as well as subscriptions costs for cloud computing services, there is a small capital risk. Though there has been much research into OR, there are areas that need to be addressed so that OR may be employed at a level to make a significant impact within manufacturing and subsequently industry practices.

Acknowledgment

This work was supported by the Science Foundation Ireland under Project:16/RC/3918. Also supported by the European Regional Development Fund and Ireland's European Structural and Investment Funds Programmes 2014-2020.

References

[1] X. Luo, D. S. Jayas, and S. J. Symons, "Identification of damaged kernels in wheat using a colour machine vision system," *Journal of Cereal Science*, vol. 30, no. 1, pp. 49–59, 1999.

[2] A. Leemans-V, Magein-H, and M.-F. Destain, "Defects segmentation on 'Golden Delicious' apples by using colour machine vision.," *Computers and Electronics in Agriculture*, vol. 20, pp. 117–130, 1998.

[3] A. O' Riordan, D. Toal, T. Newe, and G. Dooly, "Stereo Vision Sensing : Review of existing systems," in *ICST*, pp. 178–184, 2018.

[4] H. Murase and S. K. Nayar, "Learning and recognition of 3D objects from appearance," *Qualitative Vision, 1993., Proceedings of IEEE Workshop on*, vol. 24, pp. 39–50, 1993.

[5] K. Jarrett, K. Kavukcuoglu, M. Ranzato, and Y. LeCun, "What is the best multi-stage architecture for object recognition? BT - Computer Vision, 2009 IEEE 12th International Conference on," *Computer Vision, 2009*, pp. 2146–2153, 2009.

[6] K. Liu, Y. Q. Cheng, and J. Y. Yang, "Algebraic feature extraction for image recognition based on an optimal discriminant criterion," *Pattern Recognition*, vol. 26, no. 6, pp. 903–911, 1993.

[7] R. Rigamonti and E. P. F, "Learning separable filters," pp. 2754–2761, 2013.

[8] V. Martin, "Wavelets and filter banks theory design," *IEEE Transactions on Signal Processing*, vol. 40, no. 9, pp. 2207–2232, 1992.

[9] A. Ahmed, K. Yu, W. Xu, Y. Gong, and E. Xing, "Training Hierarchical Feed-Forward Visual Recognition Models Using Transfer Learning from Pseudo-Tasks," pp. 69–82, Springer, Berlin, Heidelberg, oct 2008.

[10] K. Kavukcuoglu Marc and A. Ranzato Rob Fergus Yann LeCun, "Learning Invariant Features through Topographic Filter Maps," tech. rep.

[11] J. Mutch and D. G. Lowe, "Multiclass Object Recognition with Sparse, Localized Features," tech. rep.

[12] Jason Brownlee, "Supervised and Unsupervised Machine Learning Algorithms," 2016.

[13] M. Rio and A. T. Figueiredo, "Adaptive Sparseness for Supervised Learning," tech. rep.

[14] T. Serre, L. Wolf, and T. Poggio, "Object Recognition with Features Inspired by Visual Cortex," tech. rep.

[15] M. . A. Ranzato, F.-J. Huang, Y.-L. Boureau, and Y. Lecun, "Unsupervised Learning of Invariant Feature Hierarchies with Applications to Object Recognition," tech. rep.

[16] F. Jie Huang and Y. LeCun, "Large-scale Learning with SVM and Convolutional Nets for Generic Object Categorization," tech. rep.

[17] M. Obitko and V. Jirkovsky, "Big Data Semantics in Industry 4.0," in´ *Proceedings of the 7th International Conference on Industrial Applications of Holonic and Multi-Agent Systems - Volume 9266*, pp. 217–229, 2015.

[18] J. Wan, S. Tang, Z. Shu, D. Li, S. Wang, M. Imran, and A. V. Vasilakos, "Software-Defined Industrial Internet of Things in the Context of Industry 4.0," *undefined*, 2016.

[19] M. Rußmann, M. Lorenz, P. Gerbert, M. Waldner, J. Justus, P. Engel, and M. Harnish, "Industry 4.0: The Future of Productivity and Growth in¨ Manufacturing Industries," *The Boston Consulting Group*, p. 20, 2015.

[20] F. Almada-Lobo, *Journal of Innovation Management.*, vol. 3. jan 2016.

[21] S. Raza, P. Misra, Z. He, and T. Voigt, "Building the Internet of Things with bluetooth smart," *Ad Hoc Networks*, vol. 57, pp. 19–31, mar 2017.

[22] M. Brettel, N. Friederichsen, M. Keller, and M. Rosenberg, "How Virtualization , Decentralization and Network Building Change the Manufacturing Landscape : An Industry 4 . 0," 2014.

[23] A. Botta, W. de Donato, V. Persico, and A. Pescape, "Integration of Cloud computing and Internet of Things: A survey,"´ *Future Generation Computer Systems*, vol. 56, pp. 684–700, mar 2016.

[24] A. Varghese and D. Tandur, "Wireless requirements and challenges in Industry 4.0," in *2014 International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 634–638, IEEE, nov 2014.

[25] F. Shrouf, J. Ordieres, and G. Miragliotta, "Smart factories in Industry 4.0: A review of the concept and of energy management approached in production based on the Internet of Things paradigm," in *2014 IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 697–701, IEEE, dec 2014.

[26] J. Lee and H. Kao, "Product Services Systems and Value Creation . Proceedings of the 6 th CIRP Conference on Industrial Product-Service Systems Service innovation and smart analytics for Industry 4 . 0 and big data environment," 2014.