


ORIGINAL RESEARCH

Modelling the 2D object recognition task in manufacturing context: An information-based model

Daniela Cavallo | Salvatore Digiesi  | Giorgio Mossa

Department of Mechanics, Mathematics and Management, Polytechnic University of Bari, Bari, Italy

Correspondence

Salvatore Digiesi, Department of Mechanics, Mathematics and Management, Polytechnic University of Bari, Via E. Orabona, 4, Bari, Italy.
Email: salvatore.digiesi@poliba.it

Funding information

Italian Ministry of Education, Universities and Research (MIUR), Grant/Award Number: PRIN-2017F8BB4

Abstract

In the last decades, manufacturing systems evolved to meet the high product variety required by the market. Different products can be manufactured in the mixed-model assembly lines, with an increase in the process complexity. In these production systems, the required flexibility is mainly provided by operators in the final assembly stages. Here, human errors could lead to high economic losses. A lack is observed in available research concerning a formal quantification of manufacturing complexity considering the joint effect of shape complexity and similarity in the mix variety. This paper focuses on operator decision-making in 2D object recognition tasks, since this is the most critical task performed in mixed model assembly systems. A novel model to quantify the information content in 2D object recognition task is proposed. The model is based on the Shannon's Entropy theory and considers both shape complexity and object similarities. Numerical experiments are provided, and results obtained show the effectiveness of the model in capturing the joint effect of shape complexity and similarities on the task information content. The proposed model can be adopted in a production environment for re-allocating tasks/sub-tasks to avoid the high amount of information to be processed affecting operators' performance.

KEYWORDS

2D objects recognition task, mixed model assembly line, shape complexity, shape similarity, task information content

1 | INTRODUCTION

There is an increasing demand for product variability and operational efficiency in the current manufacturing context. Consequently, industries have faced many challenges in order to shift from traditional approaches (few models with long life cycles and a small variety of attributes) to new configurations in order to meet the market requirements [1]. In the actual manufacturing environment, a new production paradigm is recognised in high product variety, short lead time, and mass customisation [2]. This production paradigm provides different ways of managing and controlling the process, contributing to the increase of the flexibility level of the industry and of its competitiveness through the product's mass customisation [3, 4]. Hence, a fundamental requirement for the assembly systems is that they must be reactive to the customers' needs

and must provide high volumes of high-quality products with a wide variety. According to Rekiek et al. [5], mixed-model assembly lines are an enabled to manage increased variety. However, these systems are characterised by higher complexity than traditional assembly lines. The complexity and its challenges are widely acknowledged. A formal quantification of manufacturing complexity is a topic not yet fully studied in depth.

In a system, complexity means something that is “difficult to understand, describe, predict or control” [6]. Moreover, production complexity is defined as the interrelations between product variants, work content, layout, tools and support tools, and work instructions [7]. Complexity is defined as ‘the state of having many different parts connected or related to each other in a complicated way’ [8] without systematically quantifying it. In manufacturing contexts, the estimation of complexity is a

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *IET Collaborative Intelligent Manufacturing* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

challenge since it is defined in a vague way and under a subjective perspective due to the lack of common and standardised measures to assess it, especially in the presence of high product variety [9].

Despite the high automation level of modern manufacturing systems, operators still play a key role and they are asked to handle and manage many different tasks, gather information, interact with many different types of technologies and work in many modes of the production [10].

In modern assembly systems, traditionally characterised by high musculoskeletal disorder risk [11], the tasks assigned to the operators tend to have a predominant cognitive content. In assembly systems, before starting a task, the operator receives a stimulus requesting him/her to select a specific part from a pool of alternative options. The choice process involves two subsequent steps. At first, the operator must recognise the received stimulus; secondly, he/she has to select the corresponding option (component). In the current work the skills level of the operator in accomplishing the assigned task is not considered, since the task difficulty is modelled as an objective demand imposed on the subject. In evaluating the complexity of choosing tasks accomplished by operators, the similarity between the chunks included in the mix variety has to be considered. Similarity impacts the choice complexity and could affect the operator and hence the overall system performance; however, it received less attention from researchers, probably due to a low perceived impact in a mass-production environment on operators' performance. Currently, where the number of options is increased due to the mass customisation, both the number of options and their similarities must be considered in evaluating the operator's choice complexity [9]. In this paper, choice complexity in 2D object recognition task is adopted to evaluate the mental workload required by a human operator in selecting a component from a number N of options (mix variety). In the model proposed, in order to measure operators' mental workload, the shape complexity of objects in the mix of alternatives as well as their similarities are considered.

Empirical and analytical studies are available on the evaluation of complexity in manufacturing context. By observing the complexity due to the variety of mix and the manufacturing performance, a negative correlation between complexity and performance is shown by Fast-Berglund et al. [12].

According to ElMaraghy et al., the complexity is related to the amount of information to be processed [13]. More information is generated by the increase of the variety which provides opportunities for the product, process or system to behave in unexpected manners [14]. Different research studies quantify the complexity according to the entropy theory [15, 16].

Object recognition task is defined as the finding and labelling of parts of an image that match objects in a scene. Sometimes, in order to reduce the task difficulty, restrictions on the shapes and semantics of the scene are adopted. These restrictions are used by operators in object recognition tasks. An object is defined by its geometric and semantic characteristics as well as by its statistical properties [17]. The objects' comparison is based on their semantic [18], topological [19]

and geometric [20] similarities. Semantic similarity is measured based on nonspatial information, such as the attributes of the objects; topological similarity is based on the relationship between spatial objects: even minimal differences may lead to different matching results; geometric similarity or matching is based on similarities in shape, often with the exclusion of the dimensional scale [21].

To perform a recognition task, one has to establish models or general descriptions of each object to be recognised. Different methods can be used in the classification of objects; the autoregressive model parameters that represent the shapes of object boundaries detected in images [22], the Fourier descriptors which have been found useful in recognition of industrial parts and characters [23], and codons which are primarily image-based descriptors in closed 2D shapes are used as methods in object classification task [24]. These methods are independent of object size and planar orientation.

Our goal is to measure the information content of the 2D object recognition task; in order to do this, a new information-based model to evaluate the 2D object recognition Task Difficulty (TD) is proposed. It integrates the objects' shape complexity and the similarities of the objects included in the mix variety. Task Difficulty allows evaluating Mental WorkLoad (MWL). MWL is a multidimensional construct involving working memory processes ranging from attention and perception to memory and decision-making [25]. MWL represents the information processing rate required to accomplish a task in a given work environment. An increase in MWL proves to negatively affect the operator's performance [26].

The remainder of this paper is organised as follows. Object recognition and shape representation are introduced in Section 2. An overview of the existing literature on the shape complexity and the similarity in 2D objects are provided in Sections 3 and 4, respectively. Section 5 presents the developed model for measuring the information content in the 2D object recognition task as a function of both shape complexity and the similarities of objects included in the mix variety. In Section 6, a numerical case study is provided. Managerial implications and discussion are provided in Sections 7 and 8, respectively. Finally, conclusions and further research are in Section 9.

2 | OBJECT RECOGNITION AND SHAPE REPRESENTATION

The object recognition task consists of recognising all objects in a given image and labelling them in classes. Humans easily recognise the objects/things in the real world, but the same action is not easy to perform by machines. The machine has to be trained by using a number of features extracted by algorithms [27]. Grouping under the same class the objects having similar features and assigning a name to the class is called classification.

By seeing colour, shape, texture, or some other features one image can be easily recognised by the human. In the same

way, the machine performs: firstly, it extracts the features of the object and then it applies the classification algorithm to label it in a particular class.

Both image segmentation and object classification give back as the output of the recognition process. All objects are separated from the background of the image by the segmentations. Subsequently, objects are matched with trained images using some classifier and/or some similarity measure. Then the object is classified based on the minimum distance of the object from a trained image [28].

Many scientific publications are focussed on object recognition and shape representation. However, most of them are IT-oriented, and investigate the issue from an artificial intelligence (computer vision) perspective, in which recognition is considered a supervised learning problem. These models are often quite far from object recognition and shape representations performed by humans [29].

In the domain of machine learning and computer vision several techniques for object classification are adopted. Concerning the object detection and recognition the most developing domain is computer vision and image processing. The applications of object classification include target recognition, video surveillance, autonomous vehicles framework, scene understanding, and others [30].

In the computer vision domain, detection and description of key points from an image are carried out by methods like SIFT, SURF or ORB that are computationally efficient. In [31], the object recognition of industrial parts is based on hierarchical classification. Results on the application of hierarchical classification show that reducing the number of instances leads to better performance, and in general it performs better than a single method like ORB, SIFT or FREAK.

Conventional zero-shot learning is proposed by Chao et al. [32]. In the work, objects from observed images as well as unobserved images are identified. The experiment was done on AwA, CUB and ImageNet datasets of images in the wild. The Authors used stacking classifiers for the classification of observed and unobserved objects. A combination scheme of colour and shape feature extraction for object recognition is proposed in [33]. The proposed approach allows recognising the object also in the presence of occlusion and cluttering. This method is applied to different datasets and proved to be more accurate and fast. In [34], the shape features are adopted to recognise objects in an image. The shape descriptor is extracted by combining three features (Geometric moments, Sobel edge detector, and Edge Histogram). The shape descriptor is compared to the stored feature vectors of the image database and the Euclidean distance metrics is used to evaluate the similarity. A framework for object detection is proposed by Wei et al. [35]. This framework works well in cluttered images, and it is based on contour shape descriptors.

Deep learning and convolutional neural networks are widely used in object classification and allow achieving excellent performance. Object recognition methods based on deep learning can be divided into three research directions: voxel-based, point-set-based, and view-based methods [36]. In the voxel-based methods the object is represented as a mesh and

consequently it is analysed by a 3D network. In the point-set-based methods the object is represented as a set of points and the cloud of points is used for prediction [37]. The view-based methods transform the 3D objects into 2D images. The view-based methods have a large amount of data and can use a mature advanced network framework. In the case of occlusion, the views from different viewpoints can complement each other's detailed features of the object and achieve excellent recognition performance [36].

The multi-class object recognition deep learning to explore convolutional neural networks is investigated by Hayat et al. [38]. The convolutional neural network is created with normalised standard initialisation and trained with a set of sample images. The results obtained by the authors are implemented in python.

In [39], the representations and the architectures for classifying 3D shapes are investigated. The Authors studied the initialisation role of the architectures for 3D shape classification by varying the training examples and using the cross-modal transfer learning. Point-based, voxel-based and multiview networks are used. The authors found that point-based networks are more robust to point position perturbations while voxel-based and multiview networks are easily fooled with the addition of imperceptible noise to the input.

2.1 | Recognition and shape representation of 2D physical objects

As illustrated in the previous sections, object recognition and shape representation are research topics deeply investigated in the IT field. There is a general convergence on wide accepted concepts and methodologies. However, less attention has been paid in the past to the case of recognition and shape representation of physical objects. In this research field, main contributions appeared in scientific literature more than 20 years ago, and the debate is still open, because of a lack of a widely accepted theory.

The perceptual decision is the basic process for object detection and recognition in humans. The perceptual decision is how sensory information guides human behaviour in the external world. This involves gathering information through the senses, evaluating and integrating it according to the subject's current goals and internal state, and using it to produce an outcome or a motor response [40]. Objects recognition by humans is rapid and seemingly effortless [41]. The human brain recognises and differentiates objects from one another, and, similarly, groups them into categories according to common features in a hierarchical fashion [42]. Human-object interactions recognition identify not only objects but also the relationships between them, providing a deeper understanding of the semantics of visual scenes than just object recognition [43] or object detection [44].

A debate on objects detection and categorisation task by humans is still open. It is not clear if firstly the object is detected and consequently categorised, or both detection and categorisation are an outcome of the same process. In [45] it is

stated that both tasks (detection and categorisation) are based on equal performance and “as soon as you know it's there, you know what it is”. Other authors state that the detection task is easier than the categorisation one [46, 47]. The subjective aspects in objects detection and categorisation task is investigated in [48, 49] through the concept of metacognition. Metacognition is a measure of subjective experience and is defined as the ability to insight into the objective correctness of a response [48]. Authors adopt metacognition for assessing subjective experience in performing detection and categorisation tasks with degraded and masked objects, while objective performance is assessed by matching. Results obtained by [48] show that metacognition (measured by fMRI scanning) is lower in the detection task than in the categorisation one since, in categorisation task, object-related information is already stored in the human brain (long-term memory).

In modern society, visual information plays a key role. In many application areas such as architectural and engineering design, fashion, journalism, advertising, entertainment, and others, images are used. Images' visual features provide a description of their content. However, more and more objects are IT-oriented, and the aim to find a tool to search for a similar object from a database is felt by researchers. In this perspective, one feature used in describing the objects' content is the shape. These shape features are named shape descriptors.

Shape descriptors are defined by [50] as the “set of numbers that are produced to describe a given shape feature” and are used to represent a simplified 2D shape. The shape descriptor represents topologically or geometrically the shape and quantifies it according to human intuition (or task-specific requirements) [50]. According to [50–52] efficient shape features must present some essential properties such as:

- Discriminative accuracy: to accurately distinguish one shape from another based on subtle differences;
- Transformation (translation, scaling, and rotation) invariance also known as pose normalisation;
- Robustness against model degeneracies/roughness;
- Uniqueness: each shape descriptor must be uniquely coupled with a unique shape;
- Performance and memory-efficient;
- Partial matching: robust against incomplete shapes;
- Insensitive to noise: small changes in the shape to lead to small changes in the shape descriptor.

In the past many shape description techniques have been developed [50]. In the last years, new techniques have been proposed. Important shape features based on shape boundary information or both boundaries and interior content are used in shape representation [53]. The most recent known developed methods in describing shape representation and description techniques are presented below.

Contour- and region-based methods are used in describing shape representation and description techniques. Each class of methods is further divided into structural and global approaches. The whole hierarchy of the classification is shown in [53]. Instead, according to [54], 2D shape descriptors are

classified into three categories: contour-based, region-based, and Hybrid 2D Shape Descriptors. In Table 1 the most common techniques for each category are provided. For more details, the reader can refer to [54].

Contour-based shape descriptors consider only the boundary and ignore the information content into the shape. As a positive aspect, these descriptors are of low computation complexity, but the noise of the image is not well handled. Both boundary and internal information are considered in region-based descriptors [55]. Unlike the contour-based, region-based shape descriptors work well against noise and shape variations.

The hybrid approach combines the Fourier Descriptors (FD) with Angular Radial Transform (ART). ART is used to extract global features, while Fourier descriptor is used to extract local ones. Results of the experiments carried out by applying FD + ART and FD + Zernike Moment Descriptors (ZMD) are in [56].

In the next sections, the shape complexity and the shape similarities evaluating methods of 2D objects are investigated (Sections 3 and 4, respectively). The attention is focussed on entropic measures [57–59].

3 | 2D OBJECTS SHAPE COMPLEXITY

In defining the concept of complexity, Li [60] stated that “The definition of complexity cannot be unique, simply because there are many different ways to quantify these difficulties. There are many different tasks concerning which the difficulties are to be quantified”. This section discusses three different methods to quantify objects' shape complexity.

In a work conducted by Rigau et al. [57], objects' shape complexity is evaluated through integral geometry and information theory tools, adopted by the authors to perform an inner shape and an outer shape complexity measure. In their measures, the authors refer to both the Shannon Entropy, quantifying the information content of a random variable [61], and to mutual information, quantifying the shared information between two probability distributions.

Shape complexity can be analysed from two different perspectives: inside and outside the object. In the first case, its

TABLE 1 2D shape descriptors classification according to [54]

Contour-based	Fourier descriptors (FD)
	Wavelet descriptors
	Curvature scale space
	Shape context
Region-based	Zernike moment descriptors (ZMD)
	Scale invariant feature transform
	Angular radial transform (ART)
Hybrid	Fourier descriptors (FD) + angular radial transformation (ART)
	Fourier descriptors (FD) + zernike moment descriptors (ZMD)

degree of structure (interdependence between its parts) is quantified. In this case, the information shared by the interior surfaces of the object is measured. A differential of the surface will be related to another differential of the surface by uniformly distributed lines that join them, that is, make them visible to each other. In the second case, its degree of interaction between the object and its circumscribing sphere is calculated using the same uniformly distributed lines [57]. These complexity measures can be used as shape descriptors in fields such as object recognition and classification; however, the shape complexity is mainly evaluated by inner shape complexity measures (e.g. [59, 62]). For more details on how to calculate the inner 2D shape complexity the reader should refer to [57].

In [58], 2D shape complexity is estimated by means of correlates of Kolmogorov complexity—entropy measures of global distance and local angle. The model developed by the authors is tested on a dataset of 2D shapes. Results show that the 2D shape complexity is related to human perception; small complexity variations are easy to distinguish in simple 2D shapes; on the contrary, complex 2D shapes are hardly distinguished by small complexity variations. The model considers the measures of global distance entropy (C_{dis}^c), local angle entropy (C_{angle}^c), perceptual smoothness (p), and the randomness measure (R). A significant correlation to the structure of the shape is observed in C_{dis}^c , C_{angle}^c , and in p while the factor R expresses the structure stability. For more details the reader should refer to [58].

The probability density function of a random variable (x) describes the statistics of x . In a continuous system, the entropy H is defined as follows:

$$H(x) = - \int_{-\infty}^{\infty} p(x) \log p(x) dx \quad (1)$$

If a 2-based logarithmic function is adopted, then the entropy is measured in units of bits. By discretising $p(x)$, Equation (1) becomes:

$$H(x) = - \sum_i p_i \log_2 p_i \quad (2)$$

The concepts of entropy described above are applied in [59] to compute the shape information for 2D planar contours through an ad hoc defined algorithm (see Figure 1). The authors obtain the shape information content from discrete samples of continuous curves. Starting from a 2D curve α , the authors discretise it in the S samples $\alpha_j = \alpha(s_j)$ with $j = 1, \dots, S$ and s_j the generic point of the curve α . Assuming a constant arc length $\Delta s = s_j - s_{j-1}$, since there is a uniform sampling along the curve, the curvature k_j is directly proportional to the turning angle θ_j between two consecutive segments ($\overline{\alpha_{j-1}\alpha_j}$ and $\overline{\alpha_j\alpha_{j+1}}$). The authors determine the probability density function of the curvature function from the θ_j estimates. At first, they assume a number M of clusters. For each cluster, a constant curvature range is considered ($\Delta\theta = (\theta_{max} - \theta_{min})/M$). Being B_i ($i = 1, \dots, M$) the number of curvature samples θ_j belonging to the i th cluster ($\theta_j \in [\theta_{min} + (i-1) \cdot \Delta\theta; \theta_{min} + i \cdot \Delta\theta]$) the cluster probability p_i becomes:

$$p_i = \frac{B_i}{S} \quad (3)$$

and the shape information of the 2D contour is obtained as:

$$H(x) = - \sum_i \frac{B_i}{S} \log_2 \frac{B_i}{S} \quad (4)$$

4 | SIMILARITIES IN 2D OBJECTS RECOGNITION

The similarity effect on the 2D objects selection task performed by an agent depends on the brain activation [63]. There is little existing research on objects similarities in the

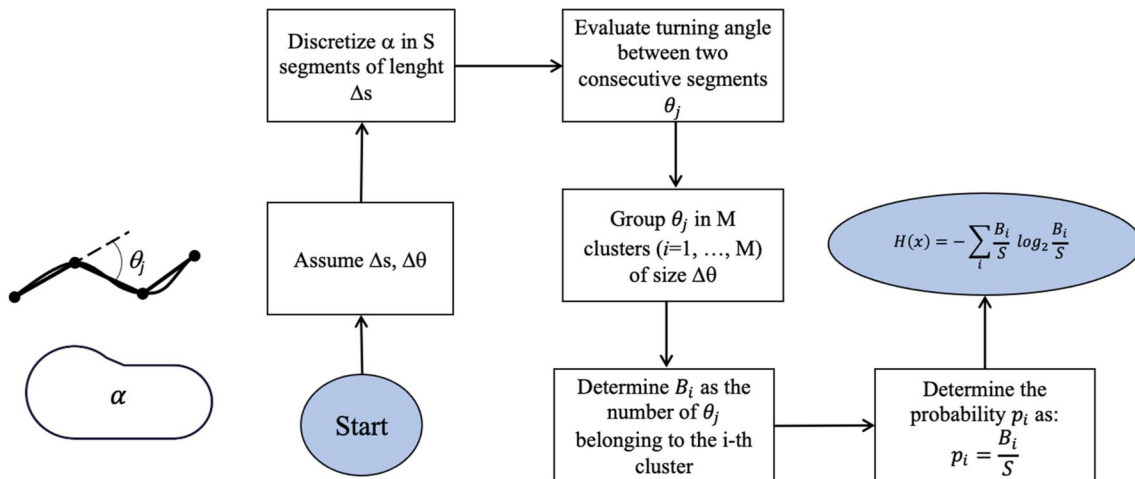


FIGURE 1 Evaluation of a 2D contour shape information as per [59]

manufacturing context, although a formal similarity measure is formalised by psychologists [64].

One of the most useful tools for the degree of similarity between objects is the similarity measure [65]. In different fields of applications such as physical anthropology, automatic classification, ecology, psychology, citation analysis, information retrieval, patterns recognition and numerical taxonomy the functions expressing the degree of similarity of items are used [66]. An important role is played by the measures of similarity/dissimilarity between the objects. The Jaccard, Dice, and cosine similarity measures in vector space for information retrieval, citation analysis, and automatic classification are the most used measures in vector space [67].

The formalisation of similarity measures is typically based on the semantic similarity; properties of objects are represented by means of their description [53]. The commonalities and differences between two semantic representations of the objects can be taken as an indicator of similarity [64]. Hence, in case of more commonalities, the higher will be the similarity and fewer will be the differences that occur.

One of the most adopted similarity measures based on the commonalities and differences between two objects is the feature-based similarity measure. In this case, the similarity between two objects A and B ($s(A, B)$) is expressed as a function (linear combination of an unstructured set of features) of their common ($A \cap B$) and distinct ($A - B$ and $B - A$) features as shown in Equation (5) [68].

$$s(A, B) = F(A \cap B, A - B, B - A) \quad (5)$$

In psychology, geometric models were initially used to exploit the analogy to space for measuring similarity [69]. Geometric models are based on the notion of multi-dimensional vector spaces. Each dimension is used to describe the properties of objects [70]. Most geometric models focus on modelling only objects.

Multi-dimensional scaling (MDS) is a dimensionality reduction technique that converts multidimensional data into a lower dimension space, while keeping the intrinsic information [71, 72]. MDS is a family of techniques for the analysis of proximity data on a set of stimuli to reveal the hidden structure underlying the data. The proximity data can be obtained from similarity judgements, identification of confusion matrices, grouping data, same-different errors or any other measure of pairwise similarity [73].

Many parallels can be drawn between geometric models and multi-dimensional scaling (MDS) models [64], but there also exist some differences: MDS uses as input the subjects' judgements about pair-wise similarities and determines the number of dimensions. Geometric models adopt a given set of dimensions and determine their values to describe each object, so that similarity can be obtained by spatial distance calculation. Geometric models are based on the analogy of semantic to spatial distance.

The similarity is a function of the spatial distance, and the most adopted similarity measure is the Minkowski distance

measure (Equation 6). In the Minkowski distance measure, stimuli are described by means of a set of dimensions collocating them as points in a multidimensional space; similarity between stimuli is inversely related to the distances of the corresponding points in the multidimensional space.

$$d_{ij} = \left[\sum_{k=1}^n |x_{ik} - x_{jk}|^r \right]^{\frac{1}{r}} \quad (6)$$

where n is the number of dimensions, x_{ik} is the value of dimension k for stimulus i and x_{jk} is the value of dimension k for stimulus j . In Minkowski measure, $r = 1$ results in the city-block distance and $r = 2$ in the Euclidian distance. The similarity is obtained as a linear (sometimes also exponentially) decaying function of the Minkowski distance d_{ij} [74, 75].

Equation (6) considers the differences in the object properties. Some properties (e.g., volume, dimension, etc.) can be easily comparable, while others (e.g., complex shape, etc.) represent a big challenge. In the last case, the measure is often based on Boolean values (i.e., true or false) as in (Equation 7).

$$|x_{ik} - x_{jk}| = \begin{cases} 0 & \text{if both } i \text{ and } j \text{ possess feature } x_k \\ 1 & \text{otherwise} \end{cases} \quad (7)$$

The semantic distance d_{ij} is then converted into the similarity measure (Equation 8) [76].

$$s(i, j) = e^{-c \times d_{ij}} \quad (8)$$

where $s(i, j)$ is the similarity measure between objects i and j , and c is the sensitivity parameter (see Figure 2).

A distance measure based on the Hausdorff distance between two points is defined in [77]. Twenty-four (24) different distance measures are compared in matching two sets of edge points extracted from two 2D shapes. The authors find that the modified Hausdorff distance (MHD) proposed performed better than the remaining 23 distance measures considered in object matching.

A framework allowing to measure the similarity of 2D objects images is developed in [78] for simple (no self-intersection) polygonal curves. To reduce the noise of 2D image digitisation, the authors propose a methodology to compare simplified shapes obtained through a digital curve evolution process.

In [79], a recognition model is proposed for comparing the couple (shape A and shape B) of 2D polygonal shapes. The authors define cost function $d(A, B)$ with the aim of measuring their dissimilarity. The cost function defined is a metric invariant to translation, rotation, and change-of-the-scale of shapes, and it is reasonably easy to compute.

In [80], a normalised measure for assessing the contour-based object pose is developed. The developed measure is applied to binary images for the assessment of known-object recognition and localisation.

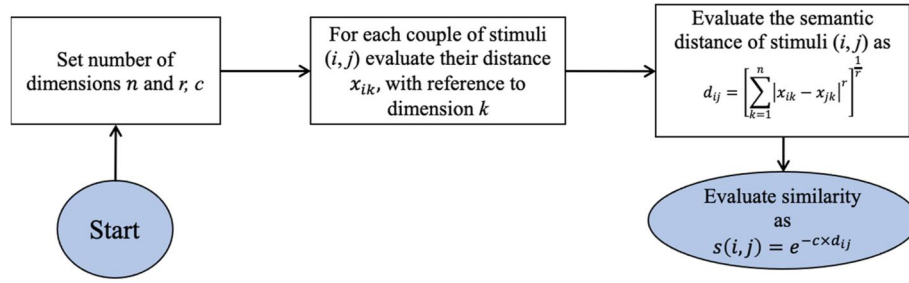


FIGURE 2 Evaluation of similarity as per [74, 75]

5 | 2D OBJECT RECOGNITION TASK MODEL

A novel model to quantify the information content in the 2D object recognition task is proposed here. The model is based on the Shannon's Entropy theory, and both shape complexity and object similarities are considered and quantified in the model. Originally used as a measure of uncertainty, the information entropy or so-called Shannon entropy, is widely adopted as a measure of complexity in many manufacturing processes.

The model proposed has been defined starting from the work of Busogi et al. [9]. The main features of the model of Busogi et al. are detailed in the following.

In a mixed-model assembly line, during the selection task, the operator receives a stimulus before he or she proceeds to select the right option. Several sequential choices such as tool choice, fixture choice and so on can be required at the workstation level. Let $k = 1, 2, \dots, K$ be the choice activities at the i th station, where K is the maximum number of choices that can be made at the i th station. Two random variables are defined: X_i^k is the outcome of the targeted variant and represents the stimulus at the k th choice task, while Y_i^k is the corresponding operator's choice. Figure 3 shows the mixed-model assembly line environment.

The same sample space is used to define both X_i^k and Y_i^k (Equation 9)

$$X_i^k, Y_i^k \in \Omega_i^k = \{v_{ij}^k | j = 1, \dots, N\} \quad (9)$$

$$H(X_i^k Y_i^k) = - \sum_{j=1}^N \sum_{t=1}^N p_{X_i^k}(v_{ij}^k) p_{(Y_i^k|X_i^k)}(v_{it}^k | v_{ij}^k) \log_2 p_{(Y_i^k|X_i^k)}(v_{it}^k | v_{ij}^k) - \sum_{i=1}^N p_{X_i^k}(v_{ij}^k) \log_2 p_{X_i^k}(v_{ij}^k) \quad (11)$$

where v_{ij}^k is the j th variant that can be selected in the k th decision-making process of the i th station, and N is the overall possible alternatives (variants) that can be selected in the k th decision-making process of the i th station.

In terms of information entropy, the successive juxtaposition of information is equivalent to the overall information

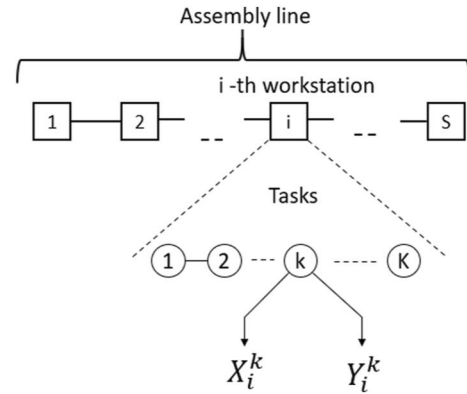


FIGURE 3 Mixed-model assembly line environment

entropy contained in variables $H(X_i^k Y_i^k)$ with joint distribution $p_{X_i^k Y_i^k}$ [9]:

$$H(X_i^k Y_i^k) = H(Y_i^k | X_i^k) + H(X_i^k) \quad (10)$$

where $H(X_i^k)$ is the average information gained by acquiring the targeted variant (per stimulus), and $H(Y_i^k | X_i^k)$ is the average information required for the selection of the part after the acquisition of the stimulus. Equation (10) is derived from the fundamental property of conditional entropy, and according to [9] the Equation (10) becomes:

$p_{(Y_i^k|X_i^k)}(v_{ij}^k | v_{it}^k)$ is the probability that the operator selects v_{ij}^k after receiving the stimulus to select it. The probability of selecting part ' j ' when ' i ' is requested is denoted $p_{(ij)}$ and based on a fuzzy logical model it can be obtained as [81]:

$$p_{(ij)} = \frac{s(i,j)}{\sum_{l \in N} s(l,j)} \quad (12)$$

where $s(i,j)$ is the similarity between part ' i ' and ' j ', and N is the set of all alternatives. Thus, based on Equation (12), Equation (11) can be extended as follows.

$$H(X_i^k Y_i^k) = - \sum_{j=1}^N \sum_{t=1}^N p_{X_i^k}(v_{ij}^k) \frac{s_{jt}}{\Theta_{v_{it}^k}} \log_2 \frac{s_{jt}}{\Theta_{v_{it}^k}} - \sum_{i=1}^N p_{X_i^k}(v_{ij}^k) \log_2 p_{X_i^k}(v_{ij}^k) \quad (13)$$

where $s_{jt} = s(v_{ij}^k, v_{it}^k)$ is the similarity between v_{it}^k (the target option) and v_{ij}^k (the j th variant), $\Theta_{v_{it}^k} = \sum_{l \in N} s(v_{it}^k, v_{il}^k)$ is the overall level of perceived similarity associated with the target variant v_{it}^k and N is the set of all available alternatives. $p_{X_i^k}(v_{ij}^k)$ is the inverse of the total number of possible alternative variants, and $\sum_j p_{X_i^k}(v_{ij}^k) = 1$.

By substituting Equation (4) in (13), the overall information entropy is obtained as follow:

$$H(X_i^k Y_i^k) = - \sum_{j=1}^N \sum_{t=1}^N \frac{s_{jt}}{\Theta_{v_{it}^k}} \log_2 \frac{s_{jt}}{\Theta_{v_{it}^k}} - \sum_{i=1}^S \left(\frac{B_i}{S}\right)^k \log_2 \left(\frac{B_i}{S}\right)^k \quad (14)$$

Equation (14) represents the overall Task Difficulty (TD).

Mental WorkLoad (MWL) is generally defined as the interaction between the operator and tasks [82]. It represents the operator information processing load while performing an assigned task [83]. Before starting a task, an operator receives a stimulus requesting him/her to select a specific part from a pool of alternative options. The choice process involves two successive steps; the operator receives a stimulus and then he/she proceeds to select the corresponding option.

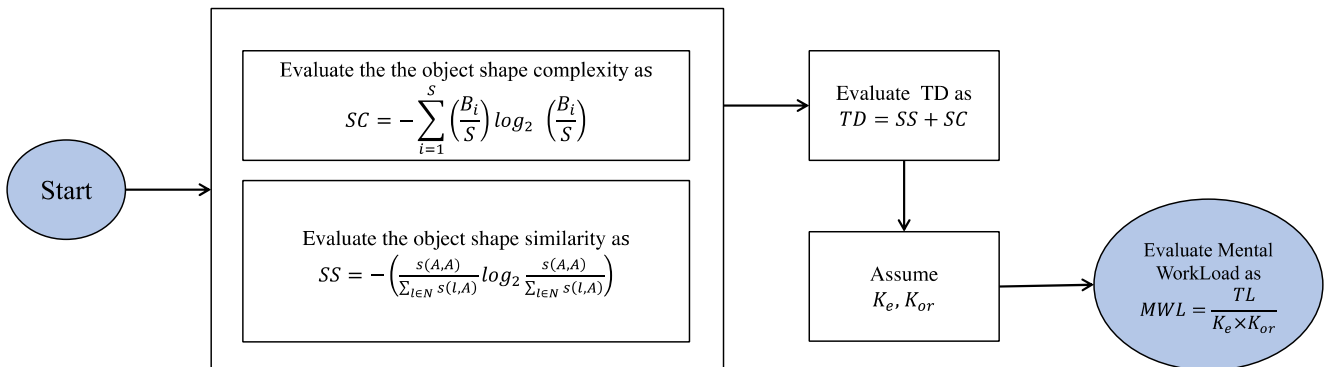


FIGURE 4 Mental workload evaluation

Starting from the Task Difficulty (TD), measured in bit units, the corresponding Task Load (TL) is obtained with reference to the given time window (T) to complete the sub-task/task, according to Equation (15).

$$TL = \frac{TD}{T} \quad (15)$$

The Task Difficulty (TD) represents the task's amount of information in selecting the right alternative from a pool and is based on both [9, 59] models. Therefore, both shape similarities and shape complexity in 2D object recognition task are considered. The TD of recognising the object "A" from a pool of alternative options (N) is obtained according to Equation (16).

$$TD = \text{Shape similarity (SS)} + \text{Shape complexity (SC)} \quad (16)$$

$$SS = - \left(\frac{s(A,A)}{\sum_{l \in N} s(l,A)} \log_2 \frac{s(A,A)}{\sum_{l \in N} s(l,A)} \right) [\text{bit}] \quad (16a)$$

$$SC = - \sum_{i=1}^S \left(\frac{B_i}{S}\right) \log_2 \left(\frac{B_i}{S}\right) [\text{bit}] \quad (16b)$$

Equation (16a) considers the similarity between the objects considered and is based on the probability of selecting object 'A' when 'A' is requested, while Equation (16b) considers the shape complexity of object A. In Equation (21a), the similarity is obtained as an exponentially decaying function of the Minkowski distance (Equation 8), and $s(A,A) = 1$.

According to Salvendy and Bü [72], the quantification of TL as well as of the environmental and organisational factors [84] allows evaluating the Mental WorkLoad (MWL) of the operator in performing the task [Equation (17), Figure 4].

$$MWL = \frac{TL}{K_e \times K_{or}} \quad (17)$$

where MWL is the Mental WorkLoad imposed on the subject and expressed in (bit/s), TL is the task load expressed in (bit/s), and K_e and K_{or} are dimensionless parameters evaluating the environmental and the organisational factors, respectively.

A scale ranging from 0 to 1 for K_e and K_{or} can be defined: in the case of $K_e = K_{or} = 1$ the system is providing a subject with satisfied stimuli and an ideal environment; if $K_e = K_{or} \rightarrow 0$ people cannot work properly.

The organisational factors are represented by job stimuli and job satisfaction. According to Judge and Klinger [85], many theories and measurements of job stimuli and job satisfaction are available. However, the two most extensively validated employee attitude survey measures are the Job Descriptive Index (JDI) [86, 87] and the Minnesota Satisfaction Questionnaire (MSQ) [88]. The JDI assesses satisfaction with five different job areas: pay, promotion, co-workers, supervision, and the work itself. This index is reliable and has an impressive array of validation evidence. The MSQ has the advantage of versatility (long and short forms are available), as well as faceted and overall measures.

The environmental factors are represented by danger, noise, temperature, and space. Acceptability of indoor environmental factors is assessed using scales ranging from ‘clearly acceptable’ (coded as 1) to ‘clearly unacceptable’ (coded as -1); the scales are presented in Standard EN 15251 [89]. The question about the acceptability of the indoor environment is formulated in the following way: “How do you assess thermal environment/air quality/sound quality/light quality/quality of the indoor environment at the moment?”.

6 | NUMERICAL CASE STUDY

In order to evaluate the MWL in the case of a 2D object recognition task, a numerical case study has been developed. The task difficulty (TD) of a specific task was evaluated,

TABLE 2 Square's side and circle diameter dimensions

Size	Square [side (cm)]	Circle [diameter (cm)]
Small	1.5	1.5
Large	3.0	3.0

TABLE 3 Shape complexity (bit)

Element considered	Calculation data	Shape complexity [SC (bit)]
Small square	$S = 16; M = 2$	0.8
	$p_1 = p(0 < \theta \leq \frac{\pi}{2}) = 4/16$	
	$p_2 = p(\frac{\pi}{2} < \theta \leq \pi) = 12/16$	
Large square	$S = 32; M = 2$	0.5
	$p_1 = p(0 < \theta \leq \frac{\pi}{2}) = 4/32$	
	$p_2 = p(\frac{\pi}{2} < \theta \leq \pi) = 28/32$	
Small circle	$S = 8; M = 1$	0.0
	$p_1 = p(\frac{\pi}{2} < \theta \leq \pi) = 8/8$	
Large circle	$S = 24; M = 1$	0.0
	$p_1 = p(\frac{\pi}{2} < \theta \leq \pi) = 24/24$	

analysing the main outcome of the Token Test ad hoc modified. The Token Test was introduced originally in 1962 as a brief test by De-Renzi and Vignolo to examine auditory comprehension deficits in aphasic patients, by having patients respond gesturally to the tester's verbal command [90]. The Token Test is a portable test that contains 20 plastic token stimuli of two sizes (large and small), two shapes (square and circle), and five colours. Here, the selection of plastic token elements simulating the 2D object recognition task is considered. The sizes of the Token Test elements are in Table 2.

Blue, black, red, green, and yellow are the five colours considered.

Four numerical cases (I to IV) were developed by varying the number of plastic token elements. The Task Difficulty (TD) of four selected plastic token elements according to Equation (16) was calculated for each numerical case. Consequently, the MWL was obtained by means of Equation (17).

By Equation (16b), the shape complexity of small/large squares and circles was calculated according to Page et al. [59]. Constant values of $\Delta S = 0.375$ (cm) and $\Delta \theta = \pi/2$ were assumed. For each element, calculation data and the amount of information related to the shape complexity are in Table 3.

In the model of Page et al. [59], the shape complexity of the circles is zero; it is considered as a reference shape.

To obtain the similarity between objects “ i ” and “ j ” [Equation (8)], the perimeter and colour were set as dimensions (Section 4). The perimeter is measured in (cm), while colour is measured in (nm) since the hue defines the colour (green, red, etc.). Hue physically depends on the dominant wavelength. The wavelengths of the visible spectrum extends from 400 to 750 nm ($1 \text{ nm} = 10^{-9} \text{ m}$) (Figure 5).

The perimeter values of small/large squares and circles are in Table 4.

The average dominant wavelength for the five colours considered is shown in Table 5.

For both dimensions (perimeter and average dominant wavelength) by means of Equation (6) the Minkowski distance measure was calculated, by assuming $\epsilon_i = r = 1$. The semantic distance obtained is then converted into the similarity measure

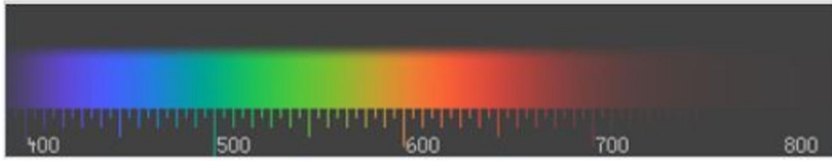


FIGURE 5 Wavelengths of the visible spectrum

TABLE 4 Perimeter values (cm)

Size	Square [perimeter (cm)]	Circle [perimeter (cm)]
Small	6.00	4.71
Large	12.00	9.42

TABLE 5 Average dominant wavelength (nm)

Colour	Average dominant wavelength (nm)
Blue	470
Black	750
Red	665
Green	525
Yellow	575

using Equation (8). In Equation (8), a unitary value of the general sensitivity parameter is adopted ($c = 1$).

Note that for N number of options, a $N \times N$ distance matrix exists. Based on Equation (8), the distance matrix can be transformed into a similarity matrix (Sim), as shown in Equation (18).

$$\text{Sim} = \begin{bmatrix} S_{AA} & \cdots & S_{AV} \\ \vdots & \ddots & \vdots \\ S_{VA} & \cdots & S_{VV} \end{bmatrix} \quad (18)$$

The overall level of similarity is the sum of the pairwise similarities between the element to be selected and every other available option, which is the sum of the row corresponding to the i th element to be selected of the similarity matrix. Next, the shape similarity amount of information (SS) between the objects included in the pool of alternatives options was calculated by Equation (16a).

The objects recognition tasks considered consist of selecting four elements from a pool of alternatives options ($N = \text{variable}$). Once the subject selects the element, it is relocated to the pool of alternatives options. So, the number of alternative options to be considered is always equal to the initial value. By using Equation (16) for each of these four elements, the Task Difficulty (TD) is obtained.

Subsequently, by means of Equation (17), in the case of selection of the required element A (numerical case I, II and IV) and element L (numerical case III), the Mental WorkLoad (MWL) in the ideal case of $K_e = K_{or} = 1$ and the Mental WorkLoad (MWL¹) in the case of $K_e = K_{or} = 0.75$ are obtained.

In Section 6.1 the four numerical cases (I to IV) where the shape and the number of alternatives options vary are presented. Results obtained are also reported.

6.1 | Results obtained

The following section shows, for each numerical case, TD, MWL, and MWL¹ values.

Numerical case I: $N = 20$; all elements considered; tasks: selecting the element-A -R -Q and -G (Table 6).

Numerical case II: $N = 10$; only squares elements considered; tasks: selecting the element-A -C -M and -Q (Table 7).

Numerical case III: $N = 10$; only circles elements considered; tasks: selecting the element-L -R -U and -G (Table 8).

Numerical case IV: $N = 10$; both squares and circles elements considered; tasks: selecting the element-A -R -M and -G (Table 9).

In a mixed-model assembly line, at the workstation level, the operators make several sequential choices. The objects' recognition tasks considered here consist in selecting in sequence the A- and R-element from a pool of alternatives options ($N = \text{variable}$). $N = 20$ (Numerical case I) and $N = 10$ (Numerical case IV) are considered as a pool of alternative options. In each numerical case study, the Overall Task Difficulty (OTD) is given by the sum of the Task Difficulty (TD) of the A- and R-element (Figure 6).

7 | DISCUSSION

This study indicates that the information-based model developed can be adopted to assess the Task Difficulty (TD) and the Mental WorkLoad (MWL) of individuals in accomplishing tasks with prevalent cognitive demand. Here, the Task Difficulty (TD) represents the task's amount of information in selecting the right alternative from a pool; both shape similarities and shape complexity in 2D object recognition task are considered. The MWL is an objective demand and depends on the environmental and organisational parameters and on the TD of the task performed by the operator. In Section 6 different numerical cases are considered where the shape and the number of alternative options vary. Four elements are selected from the pool of alternative options in each case considered. Consequently, the TD and the MWL (varying the given time windows T) have been calculated.

TABLE 6 Numerical case I

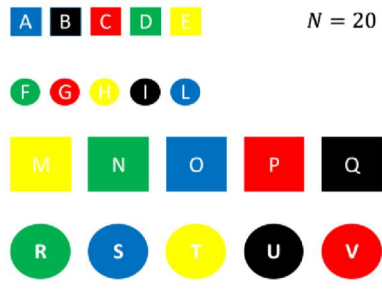
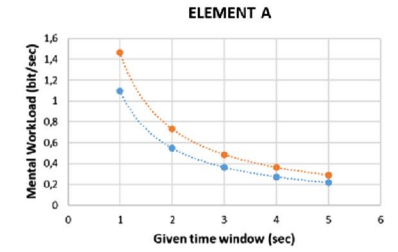
 <p>$N = 20$</p>	<table border="1"> <thead> <tr> <th>Element to be sorted</th> <th>TD (bit)</th> </tr> </thead> <tbody> <tr><td>A</td><td>1.10</td></tr> <tr><td>R</td><td>0.14</td></tr> <tr><td>Q</td><td>0.60</td></tr> <tr><td>G</td><td>0.28</td></tr> </tbody> </table>	Element to be sorted	TD (bit)	A	1.10	R	0.14	Q	0.60	G	0.28	
Element to be sorted	TD (bit)											
A	1.10											
R	0.14											
Q	0.60											
G	0.28											

TABLE 7 Numerical case II

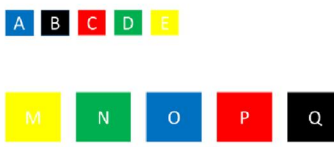
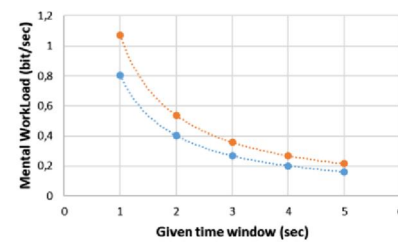
 <p>$N = 10$ square</p>	<table border="1"> <thead> <tr> <th>Element to be sorted</th> <th>TD (bit)</th> </tr> </thead> <tbody> <tr><td>A</td><td>0.81</td></tr> <tr><td>C</td><td>0.81</td></tr> <tr><td>M</td><td>0.51</td></tr> <tr><td>Q</td><td>0.51</td></tr> </tbody> </table>	Element to be sorted	TD (bit)	A	0.81	C	0.81	M	0.51	Q	0.51	
Element to be sorted	TD (bit)											
A	0.81											
C	0.81											
M	0.51											
Q	0.51											

TABLE 8 Numerical case III

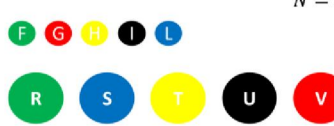
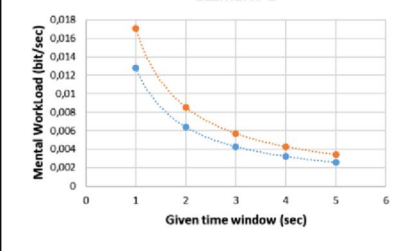

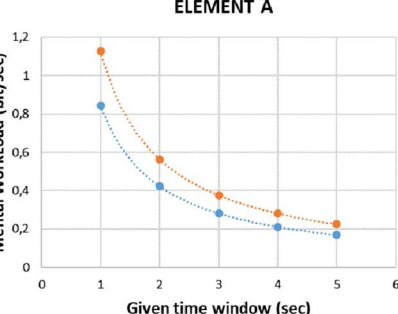
 <p>$N = 10$ circle</p>	<table border="1"> <thead> <tr> <th>Element to be sorted</th> <th>TD (bit)</th> </tr> </thead> <tbody> <tr><td>L</td><td>0.01</td></tr> <tr><td>R</td><td>0.01</td></tr> <tr><td>U</td><td>0.01</td></tr> <tr><td>G</td><td>0.01</td></tr> </tbody> </table>	Element to be sorted	TD (bit)	L	0.01	R	0.01	U	0.01	G	0.01	
Element to be sorted	TD (bit)											
L	0.01											
R	0.01											
U	0.01											
G	0.01											

TABLE 9 Numerical case IV

 <p>$N = 10$ mix</p>	<table border="1"> <thead> <tr> <th>Element to be sorted</th> <th>TD (bit)</th> </tr> </thead> <tbody> <tr><td>A</td><td>0.84</td></tr> <tr><td>R</td><td>0.11</td></tr> <tr><td>M</td><td>0.51</td></tr> <tr><td>G</td><td>0.58</td></tr> </tbody> </table>	Element to be sorted	TD (bit)	A	0.84	R	0.11	M	0.51	G	0.58	
Element to be sorted	TD (bit)											
A	0.84											
R	0.11											
M	0.51											
G	0.58											

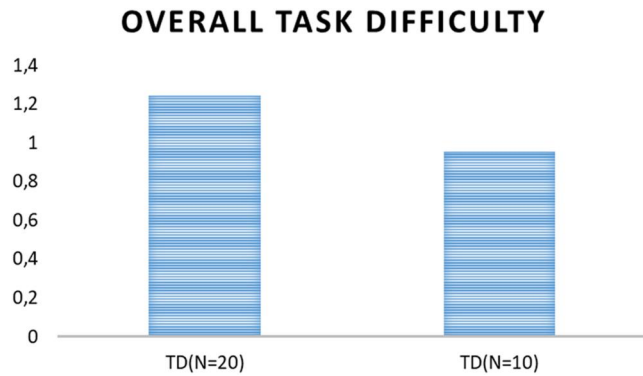


FIGURE 6 Overall Task Difficulty in case of $N = 20$ and $N = 10$ (Numerical case I and IV, respectively)

TABLE 10 A-element selection Task Difficulty (TD) in the numerical cases considered

Numerical case	TD (bit; A-element selection)
I	1.10
II	0.81
IV	0.84

By comparing values in Tables 6–9 it is noticed that the task difficulty is greater when the number of alternative options in the pool increases. Moreover, considering the numerical case III, the Task Difficulty obtained by Equation (16) is close to zero bit (Table 8) for all the four elements considered: this is due to the fact that the amount of information associated with the circular shape complexity is equal to 0 bit (since the circle is the simplest shape). Indeed, such a curve is considered having zero information. Here, only the component linked to the amount of information due to the similarities between the elements included in the pool of alternative options is taken into account.

If the task to be performed is a sequence of elements to be sorted by the operator, the Overall Task Difficulty (OTD) is given by the sum of the Task Difficulty (TD) of the task performed to select the single elements. The greater the number of alternative options, the higher will be the OTD (Figure 6).

In Table 10 the TD values in selecting A-element (small blue square) for each numerical case considered are summarised.

As shown in Table 10, the TD of the A-element selection task varies both with the number of alternative options and with the shape of the elements considered. The higher the number of alternative options available, the greater the TD. Furthermore, if the number of alternative options to be considered is the same, the TD varies according to the shape of the elements in the pool of alternatives. Indeed, the TD value is lower if elements in the pool have the same shape.

For each numerical case considered in Tables 6–9 the Mental WorkLoad (MWL) and MWL^I are represented in blue and orange colour, respectively. The MWL^I values are higher

than MWL. This means that the external stimuli (independent of the task performed) can increase the number of bits that the operator has to process in a given time window (T) to perform the same task.

Suppose the operator's mental overload grows and he/she is unable to carry out the task in the given time window. In that case, the MWL must be reduced to avoid a decrease in his/her performance and the resulting (negative) effects on the production system. In these cases, it is possible to modify:

- Task Difficulty (TD): to decrease the TD by varying the number and/or type of alternative options to be considered;
- Time window (T): by increasing T , the MWL decreases;
- K_e and K_{or} : by improving the external stimuli, the MWL decreases.

8 | MANAGERIAL IMPLICATIONS

The model proposed allows evaluating the operator's mental workload in cognitive-oriented tasks where 2D object recognition is required. In manufacturing contexts like mixed-model assembly lines, where objects' similarities and objects' shape complexity can affect operators' performance, mental workload evaluation could drive in optimising human resources to guarantee acceptable levels of error rate and safety risk. It is observed that the gap between the complexity of a task and the capabilities of the human operator could lead to an increase in human error often leading to accidents [91]. Operators' performances are significantly affected by both too high and too low mental workload [92].

Since task features (e.g., variety; similarity) affect mental workload, the re-design of critical tasks can reduce high mental workloads [93].

The model can also be applied in order to evaluate changes in operators' mental workload when the industrial setting varies due to the introduction of new products (increasing the tasks' complexity) and/or the adoption of new technologies (as in the case of I4.0 enabling technologies). In the former case, the model can be applied to verify the increase in the mental workload of critical tasks (e.g., final assembly, quality control) and the suitability of organisational strategies like job rotation in order to balance the mental workload among the operators; in the latter case, the model can be a valid theoretical support to evaluate the effectiveness (from a cognitive perspective) of technologies to be adopted.

9 | CONCLUSION AND FURTHER RESEARCH

This paper aims to assess the Task Difficulty (TD) in 2D object recognition task, and the corresponding operator's Mental Workload (MWL). Starting from literature references, a model to quantify in bit units the amount of information in

performing the recognition task (TD) considering the objects' similarities and the objects shape complexity has been proposed and applied to numerical case studies.

The case studies deal with the selection of a required element from a pool of alternative options. Different conditions by varying the number- and the shape of elements considered in the pool of alternative options have been investigated. The model allowed to quantify the TD and to evaluate the operator's MWL in the different cases. The model paves the way to industrial applications having important managerial implications, especially in the Industria 4.0 context where cognitive tasks are becoming more prevalent if compared to motor tasks. Moreover, the model could allow evaluating the optimal assignment of tasks to operators in order to meet the production targets (operators performing the given task(s) in the required time window) avoiding at the same time mental over-workload.

The model evaluations are based on numerical case studies; therefore, the reliability of the proposed model will be validated in both experimental and real industrial settings, simulating and observing 2D object recognition cognitive tasks performed by subjects differently trained and differently aged and sexed. Future research will be characterised by improvements in the understanding of how individuals' cognitive abilities affect the Mental WorkLoad, so a multidisciplinary approach, competence and empirical studies will be required. However, to get the full picture of about the cognitive support to the operators, further research and empirical studies have to be carried out due to the perceived complexity among the operators. In order to design the right cognitive support for the operator of assembly lines and to identify the optimal assignment of tasks, operators' skills and training levels have to be considered.

ACKNOWLEDGMENTS

This research was funded by the Italian Ministry of Education, Universities and Research (MIUR), SO4SIMS project (Smart Operators 4.0 based on Simulation for Industry and Manufacturing Systems-Project PRIN-2017FW8BB4).

CONFLICT OF INTEREST

Authors have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ORCID

Salvatore Digiesi  <https://orcid.org/0000-0001-8678-4771>

REFERENCES

- St. John, C.H., Cannon, A.R., Pouder, R.W.: Change drivers in the new millennium: implications for manufacturing strategy research. *J. Oper. Manag.* 19(2), 143–160 (2001)
- Koren, Y.: *The Global Manufacturing Revolution: Product-Process-Business Integration and Reconfigurable Systems* (2010)
- Yin, Y., Steckle, K.E., Li, D.: The evolution of production systems from Industry 2.0 through Industry 4.0. *Int. J. Prod. Res.* 56(1–2) (2018)
- Lucato, W.C., et al.: Model to evaluate the industry 4.0 readiness degree in industrial companies. *IFAC-PapersOnLine.* 52(13), 1808–1813 (2019)
- Rekiek, B., De Lit, P., Delchambre, A.: Designing mixed-product assembly lines. *IEEE Trans. Robot. Autom.* 16(3), 268–280 (2000)
- Sivadasan, S., et al.: Advances on measuring the operational complexity of supplier-customer systems. *Eur. J. Oper. Res.* 171(1), 208–226 (2006)
- Mattsson, S., Ekstrand, E., Tarrar, M.: Understanding disturbance handling in complex assembly: analysis of complexity index method results. *Procedia Manuf.* 25, 213–222 (2018)
- Merriam-Webster: <https://www.merriam-webster.com/dictionary/complexity> (2016)
- Busogi, M., et al.: Computational modelling of manufacturing choice complexity in a mixed-model assembly line. *Int. J. Prod. Res.* 55(20), 5976–5990 (2017)
- Mattsson, S., et al.: Forming a cognitive automation strategy for Operator 4.0 in complex assembly. *Comput. Ind. Eng.* 139, 105360 (2020)
- Intranuovo, G., et al.: Risk assessment of upper limbs repetitive movements in a fish industry. *BMC Res. Notes.* 12(1), 1–7 (2019)
- Fast-Berglund, Å., et al.: Relations between complexity, quality and cognitive automation in mixed-model assembly. *J. Manuf. Syst.* 32(3), 449–455 (2013)
- ElMaraghy, H.A., Kuzgunkaya, O., Urbanic, R.J.: Manufacturing systems configuration complexity. *CIRP Ann. Manuf. Technol.* 54(1), 445–450 (2005)
- Hu, S.J., et al.: Product variety and manufacturing complexity in assembly systems and supply chains. *CIRP Ann. Manuf. Technol.* 57(1), 45–48 (2008)
- Zhu, X.: *Modeling Product Variety Induced Manufacturing Complexity for Assembly System Design.* ProQuest Dissertation Theses (2009)
- Zhu, X., et al.: Modeling of manufacturing complexity in mixed-model assembly lines. *J. Manuf. Sci. Eng. Trans. ASME.* 130(5), 051013 (2008)
- Tieng, Q.M., Boles, W.W.: Recognition of 2D object contours using the wavelet transform zero-crossing representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 19(8), 910–916 (1997)
- Ballatore, A., Wilson, D.C., Bertolotto, M.: Computing the semantic similarity of geographic terms using volunteered lexical definitions. *Int. J. Geogr. Inf. Sci.* 27(10), 2099–2118 (2013)
- Alharthi, T.N., Elsafty, M.A.: Attribute topologies based similarity. *Cogent Math.* 3(1), 1242291 (2016)
- An, X., et al.: A shape multilevel description method and application in measuring geometry similarity of multi-scale spatial data. *Acta Geod. Cartogr. Sin.* 40(4) (2011)
- Xu, Y., et al.: Measuring the similarity between multipolygons using convex hulls and position graphs. *Int. J. Geogr. Inf. Sci.* 35(6), 1–22 (2020). <https://doi.org/10.1080/13658816.2020.1800016>
- Dubois, S.R., Glanz, F.H.: An autoregressive model approach to two-dimensional shape classification. *IEEE Trans. Pattern Anal. Mach. Intell.* 8(1), 55–66 (1986)
- Persoon, E., Fu, K.S.: Shape discrimination using fourier descriptors. *IEEE Trans. Syst. Man. Cybern. SMC.* 7(3), 170–179 (1977)
- Richards, W., Hoffman, D.D.: Codon constraints on closed 2D shapes. *Comput. Vis. Graph Image Process.* 31(3), 265–281 (1985)
- Young, M.S., et al.: State of science: mental workload in ergonomics. *Ergonomics.* 58(1), 1–17 (2015)
- Colombi, M.E., et al.: Predictive mental workload modeling for semi-autonomous system design: implications for systems of systems. *Syst. Eng.* 15(4), 448–460 (2012)
- Bansal, M., Kumar, M., Kumar, M.: 2D object recognition techniques: state-of-the-art work. *Arch. Comput. Methods Eng.* 28(3), 1147–1161 (2021)
- Chaturvedi, M., et al.: Object recognition using image segmentation. In: *PDGC 2020—2020 6th International Conference on Parallel, Distributed and Grid Computing* (2020)
- Riesenhuber, M., Poggio, T.: Models of object recognition. *Nat. Neurosci.* 3(11s), 1199–1204 (2000)
- Hussain, N., et al.: Intelligent deep learning and improved whale optimization algorithm based framework for object recognition. *Human-centric Comput. Inf. Sci.* 11, 34 (2021)

31. Merino, I., et al.: 2D features-based detector and descriptor selection system for hierarchical recognition of industrial parts. *Int. J. Artif. Intell. Appl.* 10(6), 1–13 (2019)
32. Chao, W.L., et al.: An empirical study and analysis of generalized zero-shot learning for object recognition in the wild. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 9906. LNCS (2016)
33. Diplaros, A., Gevers, T., Patras, I.: Combining color and shape information for illumination-viewpoint in invariant object recognition. *IEEE Trans. Image Process.* 15(1), 1–11 (2006)
34. Prajapati, N., Kumar Nandanwar, A., Prajapati, G.S.: Edge histogram descriptor, geometric moment and Sobel edge detector combined features based object recognition and retrieval system. *Int. J. Comput. Sci. Inf. Technol.* 7(1), 407–412 (2016)
35. Wei, H., Yang, C., Yu, Q.: Contour segment grouping for object detection. *J. Vis. Commun. Image Represent.* 48, 292–309 (2017)
36. Qi, S., et al.: Review of multi-view 3D object recognition methods based on deep learning. *Displays.* 69, 102053 (2021)
37. Yang, Z., Wang, L.: Learning relationships for multi-view 3D object recognition. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2019)
38. Hayat, S., et al.: A deep learning framework using convolutional neural network for multi-class object recognition. In: *3rd IEEE International Conference on Image, Vision and Computing (ICIVC)*, (2018)
39. Su, J.C., et al.: A deeper look at 3D shape classifiers. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11131. LNCS (2019)
40. Hauser, C.K., Salinas, E.: Perceptual decision making. In: *Encyclopedia of Computational Neuroscience*, pp. 1–21 (2014)
41. Mack, M.L., Palmeri, T.J.: The dynamics of categorization: unraveling rapid categorization. *J. Exp. Psychol. Gen.* 144(3), 551 (2015)
42. Biederman, I.: Recognition-by-components: a theory of human image understanding. In: *Human Perception: Institutional Performance and Reform in Australia* (2018)
43. He, K., et al.: Deep residual learning for image recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2016)
44. Dai, J., et al.: Object detection via region-based fully convolutional networks. In: *Advances in Neural Information Processing Systems* (2016)
45. Grill-Spector, K., Kanwisher, N.: Visual recognition: as soon as you know it is there, you know what it is. *Psychol. Sci.* 16(2), 152–160 (2005)
46. Bowers, J.S., Jones, K.W.: Detecting objects is easier than categorizing them. *Q. J. Exp. Psychol.* 61(4), 552–557 (2008)
47. Mack, M.L., Palmeri, T.J.: Decoupling object detection and categorization. *J. Exp. Psychol. Hum. Percept. Perform.* 36(5), 1067–1079 (2010)
48. Fleming, S.M., Huijgen, J., Dolan, R.J.: Prefrontal contributions to metacognition in perceptual decision making. *J. Neurosci.* 32(18), 6117–6125 (2012)
49. Meuwese, J.D.I., et al.: The subjective experience of object recognition: comparing metacognition for object detection and object categorization. *Atten. Percept. Psychophys.* 76(4), 1057–1068 (2014)
50. Yang, M., Kpalma, K., Ronsin, J.: A survey of shape feature extraction techniques. In: *Pattern Recognition Techniques. Technology and Applications*. pp. 43–90 (2008)
51. Iyer, N., et al.: Three-dimensional shape searching: state-of-the-art review and future trend. *CAD Comput. Aid. Des.* 37(5), 509–530 (2005)
52. Tangelder, J.W.H., Veltkamp, R.C.: A survey of content based 3D shape retrieval methods. *Multimed. Tool. Appl.* 39(3), 441–471 (2008)
53. Zhang, D., Lu, G.: Review of shape representation and description techniques. *Pattern Recogn.* 37(1), 1–19 (2004)
54. Kazmi, I.K., You, L., Zhang, J.J.: A survey of 2D and 3D shape descriptors. In: *Proceedings—10th International Conference Computer Graphics, Imaging, and Visualization (CGIV)*, pp. 1–10 (2013)
55. Singh, C., Pooja: Local and global features based image retrieval system using orthogonal radial moments. *Opt. Laser Eng.* 50(5), 655–667 (2012)
56. Singh, C., Pooja: An effective image retrieval using the fusion of global and local transforms based features. *Opt. Laser Technol.* 44(7), 2249–2259 (2012)
57. Rigau, J., Feixas, M., Sbert, M.: Shape complexity based on mutual information. In: *Proceedings—International Conference on Shape Modeling and Applications, SMI'05*, pp. 357–362 (2005)
58. Chen, Y., Sundaram, H.: Estimating complexity of 2D shapes. In: *IEEE 7th Workshop on Multimedia Signal Processing* (2005)
59. Page, D.L., et al.: Shape analysis algorithm based on information theory. In: *IEEE International Conference on Image Processing*, 1, 229–232 (2003)
60. Li, W.: On the relationship between complexity and entropy for Markov chains and regular languages. *Complex Syst.* 5, 381–399 (1991)
61. Shannon, C.E.: A mathematical theory of communication. *Bell Syst. Tech. J.* 27(3), 379–423 (1948)
62. Feldman, J., Singh, M.: Information along contours and object boundaries. *Psychol. Rev.* 112(1), 243 (2005)
63. Kreiman, G., Koch, C., Fried, I.: Category-specific visual responses of single neurons in the human medial temporal lobe. *Nat. Neurosci.* 3(9), 946–953 (2000)
64. Nosofsky, R.: Similarity scaling and cognitive process models. *Annu. Rev. Psychol.* 43(1), 25–53 (1992)
65. Wei, G., Gao, H.: The generalized dice similarity measures for picture fuzzy sets and their applications. *Informatics* 29(1), 107–124 (2018)
66. Ye, J.: Multicriteria decision-making method using the Dice similarity measure between expected intervals of trapezoidal fuzzy numbers. *J. Decis. Syst.* 21(4), 307–317 (2012)
67. Wei, G.: The generalized dice similarity measures for multiple attribute decision making with hesitant fuzzy linguistic information. *Econ. Res. Istraz.* 32(1), 1498–1520 (2019)
68. Tversky, A.: Features of similarity. *Psychol. Rev.* 84(4), 327–352 (1977)
69. Attneave, F.: Dimensions of similarity. *Am. J. Psychol.* 63(4), 516–556 (1950)
70. Schwering, A.: Approaches to semantic similarity measurement for geospatial data: a survey. *Trans. GIS.* 12(1), 5–29 (2008)
71. Yang, L.: Alignment of overlapping locally scaled patches for multidimensional scaling and dimensionality reduction. *IEEE Trans. Pattern Anal. Mach. Intell.* 30(3), 438–450 (2008)
72. Saeed, N., et al.: A survey on multidimensional scaling. *ACM Comput. Surv.* 51(3), 1–25 (2018)
73. Steyvers, M.: Multidimensional scaling in: *encyclopedia of cognitive science. Encycl. Cogn. Sci.* 1–5 (2002)
74. Melara, R.D., Marks, L.E., Lesko, K.E.: Optional processes in similarity judgments. *Percept. Psychophys.* 51(2), 123–133 (1992)
75. Kianimajid, A., et al.: Comparison of different methods of measuring similarity in physiologic time series. In: *IFAC-PapersOnLine*. 50(1) (2017)
76. Gallistel, R., Pashler, H.: *Stevens' Handbook of Experimental Psychology: Learning, Motivation, and Emotion*, vol. 3 (2002)
77. Dubuisson, M.-P., Jain, A.K.: A Modified Hausdorff Distance for Object Matching, pp. 566–568 (2002)
78. Latecki, L.J., Lakamper, R.: Shape similarity measure based on correspondence of visual parts. *IEEE Trans. Pattern Anal. Mach. Intell.* 22(10), 1185–1190 (2000)
79. Arkin, E.M., et al.: An efficiently computable metric for comparing polygonal shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* 13(3), 209–216 (1991)
80. Magnier, B., Moradi, B.: Shape similarity measurement for known-object localization: a new normalized assessment. *J. Imagism.* 5(10), 77 (2019)
81. Luce, R.D.: *Individual Choice Behavior: A Theoretical Analysis* (2014)
82. Paas, F., Renkl, A., Sweller, J.: Cognitive load theory and instructional design: recent developments. *Educ. Psychol.* 38(1), 1–4 (2003)
83. Sheridan, T.B., Stassen, H.G.: Definitions, models and measures of human workload. *Mental Workload*. pp. 219–233 (1979)

84. Bi, S., Salvendy, G.: A proposed methodology for the prediction of mental workload, based on engineering system parameters. *Work Stress*. 8(4), 355–371 (1994)
85. Judge, T.A., Klinger, R.: *The Science of subjective well-being (Job Satisfaction well-being at work)*, vol. 45(10) (2008)
86. Smith, P.C., Kendall, L.M., Hulin, C.L.: *The Measurement of Satisfaction in Work and Retirement: A Strategy for the Study of Attitudes* (1969)
87. Yeager, S.J.: Dimensionality of the job descriptive index. *Acad. Manag. J.* 24(1), 205–212 (1981)
88. Weiss, D.J., et al.: *Manual for the Minnesota satisfaction Questionnaire. Manual for the Minnesota Satisfaction Survey*. 125 (1967)
89. Brussels: European Committee for Standardization. EN15251. *Indoor Environmental Input Parameters for Design and Assessment of Energy Performance of Buildings Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics* (2007)
90. Spreen, O., Risser, A.H.: Assessment of aphasia. In: *Acquired Aphasia*, pp. 71–156 (1998)
91. Gomes, J.O., et al.: Ergonomics, safety, and resilience in the helicopter offshore transportation system of Campos Basin. *Work*. 51(3), 513–535 (2015)
92. Aricò, P., et al.: A passive brain–computer interface application for the mental workload assessment on professional air traffic controllers during realistic air traffic control tasks. *Prog. Brain Res.* 228, 295–328 (2016)
93. Galy, E., Cariou, M., Mélan, C.: What is the relationship between mental workload factors and cognitive load types? *Int. J. Psychophysiol.* 83(3), 269–275 (2012)

How to cite this article: Cavallo, D., Digiesi, S., Mossa, G.: Modelling the 2D object recognition task in manufacturing context: An information-based model. *IET Collab. Intell. Manuf.* 4(2), 139–153 (2022). <https://doi.org/10.1049/cim2.12048>