Overview of techniques for detecting object's features and embedding them in multidimensional spaces

R. WASILUK

rafal.wasiluk@student.wat.edu.pl

Military University of Technology, Doctoral School
Kaliskiego Str. 2, 00-908 Warsaw, Poland

It is evident that each object in the real world possesses unique properties. A subset of these characteristics can be readily described in quantitative terms. Examples of such features include the number of wheels in a vehicle, the floor area of a residential property, or the year of construction of a building. However, certain characteristics of objects exhibit a higher level of complexity. Examples of such features include object shape, color, and texture. These characteristics, frequently defined in terms of objects depicted in images, represent the primary characteristics that can be identified in real-world objects. The processing of these visual attributes has been the subject of scientific research for decades, and the literature on this topic is extensive. The objective of this article is to synthesize the existing methods for detecting object's shape, color, and texture and embedding them in multidimensional spaces. By applying these methods, it is possible to represent the features of the object as points in multidimensional spaces. Such representations can be used to solve multicriteria optimization problems.

Keywords: visual features, feature extraction, feature embedding.

DOI: 10.5604/01.3001.0055.2089

1. Introduction

A fundamental component of any decision-making process is the comparison of objects. The efficacy of this process depends on how the objects to be compared are described. The process of characterizing involves determining the values that describe each of its features. The identification of object characteristics is typically achieved through the measurement or application of various feature detection algorithms. Subsequently, the decision-maker defines its preferences regarding the relevance of specific characteristics of the objects to be contrasted. This step can be efficiently performed if the objects being compared have features that can be easily represented as a number. Examples of such facilities include houses and apartments. In their cases, the area of the apartment, the number of bedrooms, and the year of construction can be readily determined. However, in instances where the object being contrasted lacks readily distinguishable features or has been presented in the form of an image, the determination of its value becomes a more challenging task. In such cases, the detection of object features can be achieved through automatic recognition. This scientific discipline focuses on observing real-world objects and then assigning them to a class in feature space [10]. Thus, after the object is recognized, it is embedded in a multidimensional space. The specimens described in this manner can be further compared using distance measures and similarity functions. It is noteworthy that some of the existing methods encode object features explicitly and some do so implicitly. The first of these approaches gives the decision maker an additional opportunity to assign weights to a subset of the features. This approach could be advantageous in emphasizing the importance of certain characteristics. Consequently, the identification and description of object attributes emerge as important tasks in the context of comparing real-world objects. Therefore, this article will review some of the existing methods that are used in the first steps of automatic recognition – the detection of object's features and embedding them in multidimensional spaces.

2. Object's shape

A notable example of a feature that is difficult to operate with in its unprocessed form is the shape of an object. Despite its inherent complexity, this characteristic is worth considering because it is one of the most critical factors that allows humans and machines to perceive, recognize, determine pose, and so on.

The literature on shape recognition is very extensive. Consequently, this section of the article is limited to a presentation and thorough description of a few methods. Due to the multitude of existing approaches there are also different ways to categorize these techniques. One of them classifies them into two main groups: boundary-based and region-based [7]. This division is based on the operating domain of the algorithm. However, a more detailed analysis reveals a breakdown of shape recognition techniques into template matching, syntactic or structural matching, statistical classification, and artificial neural networks [1]. Furthermore, cognitive psychology has presented an alternative classification of pattern recognition techniques based on human perception of the world. The classification includes template-based, prototype-based, and feature-based approaches [4]. Despite the utilization of different mathematical algorithms and theories, all methods share the same goal – to obtain a description of a shape presented in the scene. In recent years, wavelet descriptors, curvature scale space, moment, and Fourier descriptors have dominated the field of object detection and feature embedding [6], [7]. Nowadays, due to the rapid development and high interest in artificial intelligence methods, users are more likely to use tools such as deep neural networks and long short-term memory (LSTM) networks. The majority of these techniques are focused on producing an object shape representation as a single vector with a fixed predefined shape [11], [12]. However, alternative methods exist that describe an object's shape as a set of values [5].

One of the earliest approaches to detect an object's shape with the use of a computer was a method called template matching. This method matches the pattern to the image that is likely to contain it. It is important to note that numerous methods exist for achieving this objective. One such approach, which can be considered naive, shifts the template all over the image. It searches for a location, where the similarity score between the template and the "covered" fragment of the image is the highest. However, this method is looking for an exact match to the pattern presented in the template, which may not always be the most effective approach. In addition, it does not handle situations where the target object is rotated, scaled, occluded, or the light illuminating the object is different [1]. Nevertheless, researchers specializing in template matching methods have developed some improvements to make their approach more robust and invariant to object transformations and environmental changes [2], [3]. Despite the advances in the field, the methods belonging to the template matching algorithms are computationally demanding due to their complexity. Consequently, their implementation in real-time systems is impractical [9].

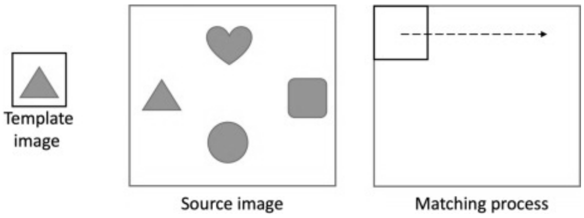

Fig. 1. A trivial example of template matching. This picture was presented in Chapter 22 of *Medical Image Analysis* book written by Cheng Lu et al.

Another approach that has been extensively utilized in the domains of shape detection and description is founded on the implementation of Fourier transforms or wavelet transforms [5], [6], [7], [8]. One of the existing methods constructs a set of values called Fourier descriptors, that represent the shape of the object. To acquire these values, the input image containing the object of interest is first preprocessed. The purpose of this step is to obtain an outline of the object (for contour-based approaches) or an area of the input image that contains the desired object (for region--based methods). This step is composed of two processes: binarization and denoising. Binarization involves converting the image to a binary representation, such that true pixels contain the object of interest, while false pixels do not. Denoising involves the removal of possible noise, which is a potential side effect of the binarization process. To represent the extracted shape in a numerical form the following methods can be used: Zernike polynomials, Legendre moments, generic Fourier descriptors, and wavelet descriptors. It is important to note that region-based approaches are usually translation and scaling invariant. This property is achieved by performing a normalization process before extracting the shape features. Nevertheless, this does not address another important issue. The proposed method is not invariant to rotation. To overcome this problem, scientists often represent the shape using the polar coordinate system [7]. The method employed in contour--based methods is distinct. If this approach is chosen, then the acquired shape of an object must be traced. The objective of this operation is to transform the filled area of the object into a contour. Based on the results from the previous step, functions that represent the contour are constructed. These representations called shape signatures may include complex coordinates, centroid distance, curvature signature, cumulative angular function and others [6], [7]. The Fourier transforms of these values are called shape descriptors and they are used to represent the contour of an object contour. Typically, the number of Fourier descriptors generated for a single object is quite large. However, not all of them are necessary to describe the desired shape with the required accuracy. Transformations with higher frequencies can describe smaller details of the object contour and are not important when considering the whole shape. Consequently, most of these high-frequency components can be disregarded. Conversely, descriptors with the lowest frequencies describe the most general features of the shape. They are the most important in the context of tasks that describe this attribute. The calculated values of these descriptors can be successfully used to retrieve images containing similar objects (for example in CBIR systems). The acquisition of images, like the input picture, is possible by comparing corresponding shape descriptor values generated for different input graphics. Contour-based shape descriptors can also be successfully used to reconstruct the contour of object when necessary [6]. One of the main advantages of this approach is invariance to shape rotation, translation, and scale [6], [7], [8]. However, region-based methods cannot be used for every case, especially when the shape of the object is complex (for example it consists of many disjoint areas).

Another approach that has been based on the use of Fourier descriptors and wavelet transforms are multiscale feature descriptors. These are generated by implementing an additional step in the process of describing the shape feature. This operation consists of applying a wavelet transform to the object's boundary function. The result of this procedure is the creation of multiple wavelet coefficients. However, these values cannot be used directly to compare the shapes of different objects. This limitation stems from the fact that the values of the wavelet coefficients are dependent on the starting point and the length of the object's boundary. To address this issue, the Fourier transformation is applied to these values. In order to achieve rotation and scale invariance the feature vectors are devoid of phase parts, and they are normalized, respectively [7], [8].

Additionally, pattern recognition and encoding methods are employed that are based on the use of Hidden Markov Models (HMM). A solution based on this approach was presented by Janusz Bobulski in his book [9]. Traditional HMMs have the capability to process one-dimensional signals. Therefore, using this mathematical tool with graphical data was not a straightforward attempt. The images intended for processing by Markov's models were required to be transformed into a vector form first. This modification resulted in the loss of spatial information encoded within the input picture. Furthermore, additional operations performed to reshape the graphical data caused computational overhead. To address these challenges, the author proposed the use of a novel structure of HMMs. These models were designed to process two-dimensional or three-dimensional data. The pattern recognition approach based on this method involves creating a separate model for each possible template. The process of creating such structures is called learning or training. This task utilizes the data that was extracted from the image of the pattern. Before the input data is fed to the model, it has to be processed through a feature extraction step. This operation was implemented to enhance the model's pattern recognition performance and accuracy. The execution of this operation can be achieved through a variety of methods, including Fourier transform, wavelet transform, and others. In this approach the HMM that obtains the highest probability of generating a provided sample best describes the given template [9], [10].

Nowadays, the most popular approaches to recognize and embed the shape of an object are using artificial neural networks. For a simple object, its outline can be recognized by the pre-trained machine learning model, which is able to identify a predefined set of form classes on which it has been trained [13]. In this method the information about a retrieved object's shape could be encoded as for example: a single number or vector of values 1 and 0 (as in the one-hot encoding method). The approach described above has a significant drawback. It can only recognize and describe a limited set of shapes. However, given the unpredictability of real-world objects, this approach is inadequate for addressing the full complexity of shape variability. In order to address this issue, researchers have developed a variety of methods for detecting and encoding an object's boundary over the past few years.

One of the earliest approaches to detect and encode object's shape using artificial neural networks was developed by a team of researchers from AT & T Bell Laboratories. In 1994, a group under the direction of Jane Bromley proposed to verify the authenticity of handwritten signatures with use of a novel artificial neural network architecture. This innovative structure was named the Siamese network by its inventors. The configuration comprised two distinct yet analogous model structures that shared the same set of weights.

As illustrated in Figure 2, the signatures after a step of preprocessing were passed through the neural networks. These structures were tasked with producing vectors of a predefined shape that contained the encoded boundaries of the handwritten signatures. The generation of these embeddings allowed the researchers to perform a straightforward signature comparison. This was accomplished by calculating the distance between the mappings of a stored ground-truth signature and a suspicious one [11]. Researchers at AT & T Bell Laboratories demonstrated that neural networks can be effectively utilized for the comparison of objects' shapes and the generation of an object's shape representation.

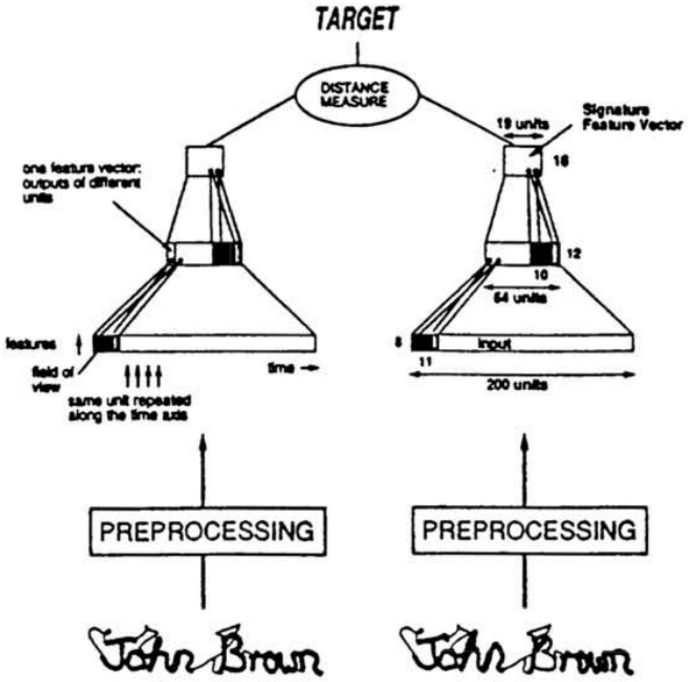

Fig. 2. The architecture of the Siamese neural network proposed by Jane Bromley et al. in [11]

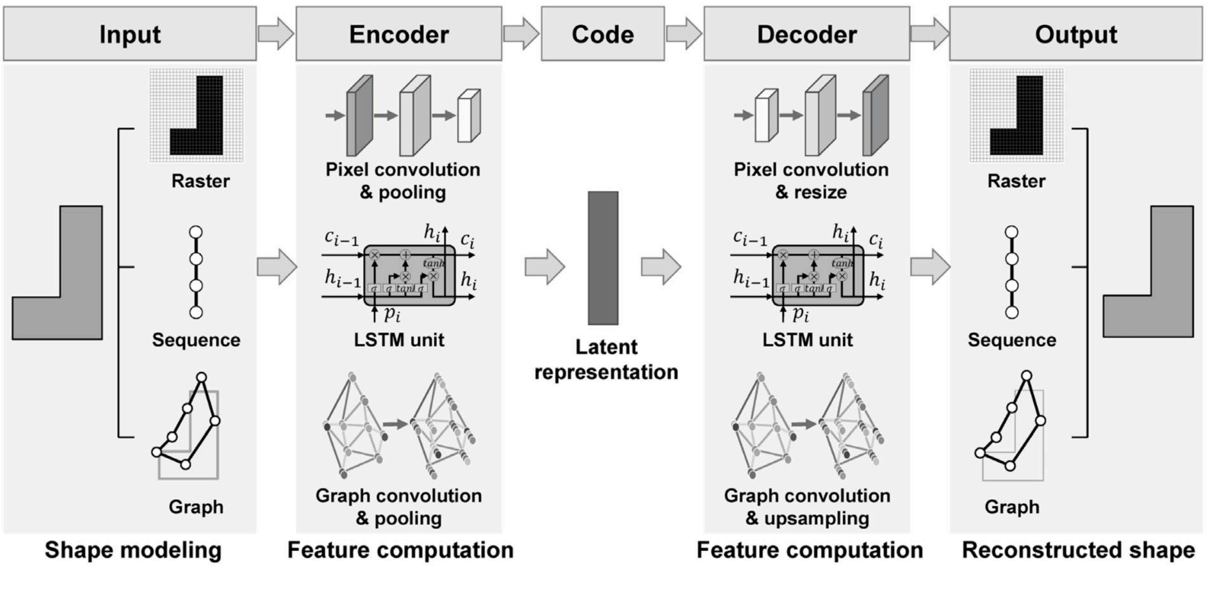

Fig. 3. Three approaches to perform building shape embedding with use of encoder-decoder architecture [12]

Further advances in artificial intelligence and deep learning domains have allowed researchers to use more sophisticated and accurate methods for shape detection and coding. In 2022, two scientists from China have published a paper proposing three different approaches to transform spatial objects into a vector representation: pixel-based, sequence-based, and graph-based [12].

As illustrated in Figure 3, all presented solutions use encoder-decoder neural network architecture for the purpose of generating embeddings of the shapes of geospatial objects. It is important to note that the solutions described by the authors were designed as an unsupervised algorithms. Consequently, the data which was fed into the model did not require any special labeling. During the training process of the first proposed approach, the model transformed the object shape into an image with the dimensions of 28x28 at the very beginning. Constructed raster deliberately rejected information about object texture and color, focusing exclusively on shape. Subsequently, the convolutional neural network encoder transformed the input into a 128--dimensional vector. The decoder was then tasked with recreating the input shape with optimal fidelity, based on the results obtained from the encoder. The sequence-based approach differed in the type of the input (a counterclockwise sequence of points constructed on a base of the original shape) and the use of LSTM cells in both the encoder and decoder instances. The final solution proposed by the authors of the article was based the concept of transforming the shape into a graph. Subsequently, the graph pooling encoder was employed to transform the resulting representation into a 128-dimensional vector. As in two previous approaches, the obtained result was reconstructed with the use of graph upsampling technique by the decoder. The experiments conducted by the researchers indicated that their solutions are sensitive to the shape rotation and change of the initial sequence point. However, the embedding produced by the pretrained encoder could be easily used as a shape descriptor.

3. Object's colors

Another feature of an object that is critical to the recognition process is the set of colors it possesses. This attribute can be defined for each pixel that represents the object, which makes it easy to recognize. However, it should be noted that, in contrast to the shape of an object, this characteristic is dependent on environmental conditions. This is because the perception of color can change depending on the intensity and temperature of the light that illuminates the object. Furthermore, the observation of this feature is even more complicated when the surface of the object is not flat and has high reflectivity. Consequently, the perception of color becomes illusory and dependent on the color of nearby objects. The subsequent section of this article will concentrate on color detection algorithms and approaches used by researchers to achieve color constancy.

One of the most simple and straightforward methods to represent color characteristics is a histogram [14], [15]. This method involves counting all the pixels that are represented by the same value. To illustrate, for the RGB model, a histogram is created for each color channel. Consequently, the overall color identity of an object is represented as three vectors of length 256, with each vector describing the red, green or blue intensities. The color representations of a given object can be further compared with the color embedding of other objects using distance measures or similarity functions. A histogram and its variations have proven to be a useful tool for the description of the colors of objects in the same environmental conditions. The simplicity of this method is one of its most significant advantages. The utilization of this approach also provides rotation and scale invariance. However, this method has some significant limitations. It generates large number of features to represent the colors of an object, and it is sensitive to environmental changes. To address this challenge, researchers usually convert the input image to grayscale and then perform color feature analysis [16].

Another approach which is widely used to describe colors of the object is color moments [17]. This method is strongly related to the method presented in the previous paragraph. The prerequisite of this method is to interpret the color distribution of an object as a probability distribution. On this basis, the values called color moments are constructed. A common approach is to store the first (1), the second (2) and the third moment (3) for each color channel available. These values respectively mean: the mean color (E_i), the variance of color (σ_i) and the skewness (s_i). Below the mathematical formulas for these characteristics are given and where: N is the number of image pixels, i is the index of color channel, j is the index of pixel, p_{ij} is the value of j pixel at the channel i.

$$E_i = \frac{1}{N}\sum_{j=0}^{N} p_{ij} \quad (1)$$

$$\sigma_i = \left(\frac{1}{N}\sum_{j=0}^{N}\left(p_{ij} - E_i\right)^2\right)^{\frac{1}{2}} \quad (2)$$

$$s_i = \left(\frac{1}{N}\sum_{j=0}^{N}\left(p_{ij} - E_i\right)^3\right)^{\frac{1}{3}} \quad (3)$$

Probability theory says that every distribution is uniquely characterized by its moments. Thus, the values computed for the color distribution of the object uniquely describe the set of colors the object has. The main advantage of this approach is that it requires much less memory to encode the feature than the histogram method.

Acquiring colors of an object in a constant and repetitive manner is a very challenging task. Even a small change in the intensity of the illumination or the use of a lower camera sensor sensitivity can cause a significant shift in the histogram representation. Therefore, the effectiveness of the methods presented earlier in this section of the article may be compromised. Consequently, in 2022 researchers led by Kim Bomi proposed a method to achieve color constancy under the variable lighting conditions. This approach is referred to as Channel-aware Color histogram Matching method for low-light image enhancement (CCM). The operation of their method consists of using a channel-aware residual neural network and a differentiable intensity histogram to capture color features [18]. The effect of restoring low-level light image with use of CCM method was illustrated in Figure 4.

4. Object's texture

The third object feature that is easy to detect by humans is object texture. Having information about this characteristic can significantly affect the perception of the object. However, recognizing it and describing it in a form that can be further used can be a challenging task. This is due to the fact that, in contrast to the object's color, this characteristic is not defined for every pixel presented in the image. It is rather specified for specific regions what makes it another spatial defined object's feature. Moreover, the patterns under consideration can have varying degrees of complexity and detail. The correct identification of them can have a major impact on the accuracy of the description process. Consequently, the fields of texture and material recognition remain in a state of open development, and researchers are constantly searching for more efficient ways to detect and represent the texture of materials.

One of the earliest approaches to describing the texture of an object was the Gray-Level Co-Occurrence Matrices (GLCM), which was introduced in 1973 [14], [19]. The objective of this method was to capture the spatial nature of the texture pattern and encode it as a numerical feature. This objective was accomplished by constructing measures – angular nearest-neighbor gray-tone spatial-dependence matrices. The design of these data structures was intended to encode information about pixel-level tonal variations in a given direction and radius. Conventionally, this approach defines four measures of this type: horizontal changes (0 degrees), vertical changes (90 degrees) and diagonal changes (45 degrees and 135 degrees). However, it should be noted that these matrices could not be used directly as texture feature descriptors. Consequently, a series of number-based features are extracted for each direction, based on the values obtained from these matrices. Examples of these include contrast, correlation, energy, and entropy.

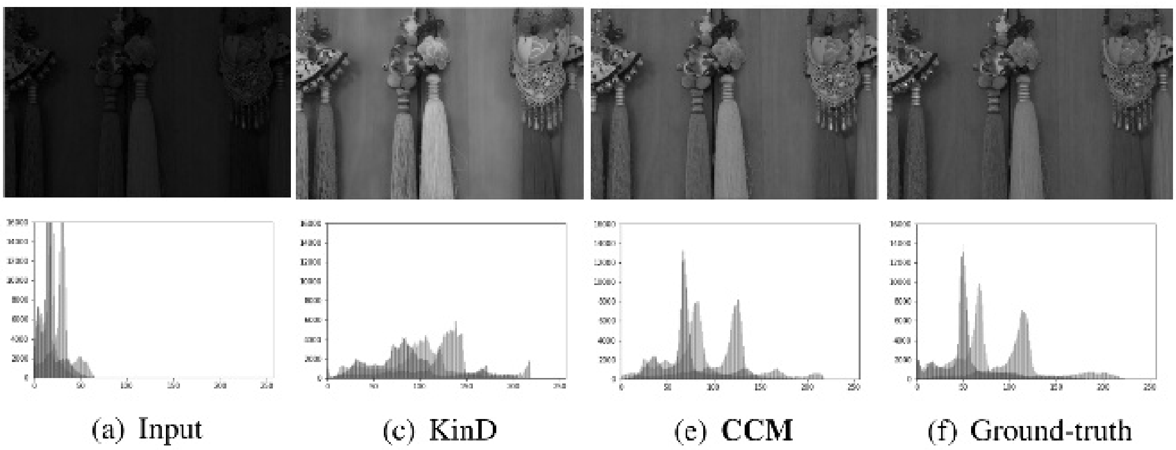

Fig. 4. The results of the process of histogram restoration using the CCM approach presented by Kim Bomi et al. in [18]

Despite the age of the solution, its modifications continue to be used in various solutions, such as the detection of melanoma in the medical field [20] or the classification of use-wear on flint tools in archaeological research [21]. It is important to note that the GLCM approach is only applicable to grayscale images. This inherent limitation naturally results in the loss of color information, which can be critical for accurate texture recognition. To address this limitation, a method known as Integrative Co-Occurrence Matrices (ICM) was proposed. This method is capable of processing objects that are represented as colorful images. Achieving this objective involves the calculation of co-occurrence matrices for each color channel individually, as well as between pairs of colors [14].

In recent years, a prevalent approach for detecting and encoding a texture was Local Binary Patterns (LBP). This method is characterized by a limited number of steps that facilitate the calculation of the value of this measure $LBP_{P,R}$ for a selected grayscale image. In the initial phase, the P equally spaced pixel values situated on a circle with defined radius R from the selected pixel are thresholded by value at this central point. The result of this step is a set of binary values, which are defined as 1 if the value of the pixel at the circle exceeds the threshold, or 0 if it does not. Subsequently, the resulting values are multiplied by consecutive powers of two and then summarized. The result of this process is called LBP measure. In the final step, the calculated LBP values are collected into a histogram. This facilitates the subsequent comparison of different pattern representations using distance measures or similarity functions [22]. The calculation of LBP for an image is illustrated in Figure 5. This method has gained significant recognition because of its advantages. The most important of these is its efficiency and simplicity. Another notable advantage of this approach is its ability to maintain consistency in the face of variations in lighting conditions, ensuring reliable results under different lighting environments. Consequently, approaches based on this method are still in use nowadays in various areas [14], [24].

The domain of image processing also knows some methods that are able to detect and describe the texture of objects using artificial intelligence tools. One such method is Multer (Multi-Level Texture Encoding and Representation), which was presented by a group of researchers in 2019. They stated that by fusing low-level and high-level feature encodings, the quality of texture identification can be significantly improved. To validate their hypothesis, they have developed a solution that utilizes deep convolutional neural networks to identify the material structure of an object. Their approach involved generating input image representations by passing them through a series of convolutional layers collected in "blocks". Subsequently, the feature vectors were passed to the LEM unit (learnable encoding module). This specialized component has the capability to transform the model block output into a fixed-size, multi-level texture descriptor. The resulting vectors, which were produced at different model levels, were subsequently merged into a single vector and fed into the classification layer [23].

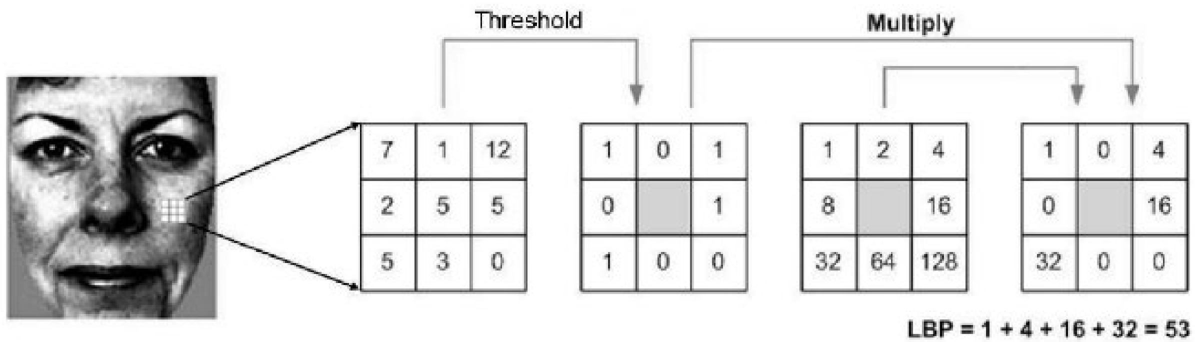

Fig. 5. An example of Local Binary Patterns calculation presented in [22]

5. Conclusions

This paper presents a compendium of methods for the detection and embedding of objects' features into a format comprehensible by mathematical algorithms. As demonstrated by the provided examples, the representations generated by these techniques can be effectively utilized to compactly represent the features of complex objects. These representations facilitate the inclusion of real-world objects in the decision space of multicriteria optimization problems. Moreover, some of the presented methods encode object's features explicitly. This allows their representation values to be purposefully manipulated to achieve a desired effect. Consequently, this facilitates the modeling of a decision-maker's preferences for these sophisticated attributes in multicriteria optimization problems.

6. Bibliography

[1] Jain A.K., Duin R.P.W., Mao J., "Statistical pattern recognition: a review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 1, 4–37, (2000).

[2] Talmi I., Mechrez R., Zelnik-Manor L., "Template Matching with Deformable Diversity Similarity", *2017 IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Honolulu, USA, 2017, p. 1311–1319.

[3] Dai W., Chang T., Zhang, L., "Template matching based on deformation diversity similarity and scale filtering", *2020 IEEE 3rd International Conference of Safe Production and Informatization* (*IICSPI*), Chongqing City, China, 2020, p. 704–708.

[4] Shugen W., "Framework of pattern recognition model based on the cognitive psychology", *Geo-Spatial Information Science*, Vol. 5, No. 2, 74–78 (2002).

[5] Zhanwei Y., Fuguo L., Peng Z., Bo C., "Description of shape characteristics through Fourier and wavelet analysis", *Chinese Journal of Aeronautics*, Vol. 27, No. 1, 160–168 (2014).

[6] Zhang D., Lu G., "Study and evaluation of different Fourier methods for image retrieval", *Image and Vision Computing*, Vol. 23, No. 1, 33–49 (2005).

[7] Yadav R.B., Nishchal N.K., Gupta A.K., Rastogi V.K., "Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: A comparative study", *Optics and Lasers in Engineering*, Vol. 45, No. 6, 695–708 (2007).

[8] Kunttu I., Lepistö L., Rauhamaa J., Visa A., "Multiscale Fourier Descriptor for Shape-Based Image Retrieval", *Proceedings of the 17th International Conference on Pattern Recognition* (*ICPR'04*), Vol. 2, p. 765–768, IEEE Computer Society 2004.

[9] Bobulski J., *Rozpoznawanie wzorców 2D i 3D na obrazach cyfrowych za pomocą ukrytych modeli Markowa*, Akademicka Oficyna Wydawnicza EXIT, Warszawa, 2018.

[10] Kwiatkowski W., *Metody automatycznego rozpoznawania wzorców*, BEL Studio, Warszawa 2007.

[11] Bromley J., Guyon I., LeCun Y., Säckinger E., Shah R., "Signature Verification using a 'Siamese' Time Delay Neural Network", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 7(4) 669–688 (1993), DOI: 10.1142/S0218001493000339.

[12] Xiongfeng Y., Min Y., "A Comparative Study of Various Deep Learning Approaches to Shape Encoding of Planar Geospatial Objects", *International Journal of Geo-Information*, Vol. 11(10), art. 527 (2022), DOI: 10.3390/ijgi11100527.

[13] Gaupta L., Dayeh M.R., Tammana R., "A neural network approach to robust shape classification", *Pattern Recognition*, Vol. 23, No. 6, 563–568 (1990).

[14] Bianconi F., Fernández A., Smeraldi F., Pascoletti G., "Colour and Texture Descriptors for Visual Recognition: A Historical Overview", *Journal of Imaging*, Vol. 7, No. 11, art. 245 (2021).

[15] Liang P., Blasch E., Ling H., "Encoding Color Information for Visual Tracking: Algorithms and Benchmark", *IEEE Transactions on Image Processing*, Vol. 24, No. 12, 5630–5644 (2015).

[16] Cusano C., Napoletano P., Schettini R., „Evaluating color texture descriptors under large variations of controlled lighting conditions", *Journal of the Optical Society of America A*, Vol. 33, No. 1, 17–30 (2015).

[17] Stricker M.A., Orengo M., "Similarity of color images", *IS & T/SPIE Symp. on Electronic Imaging: Science & Technology*, p. 381–392, Int. Soc. for Optics and Photonics, 1995.

[18] Bomi K., Sunhyeok L., Nahyun K., Donggon J., Dae-Shik K., "Learning Color Representations for Low-Light Image

Enhancement", *2022 IEEE/CVF Winter Conference on Applications of Computer Vision* (*WACV*), p. 904–912, Waikoloa, USA, 2022.

[19] Haralick R.M., Shanmugam K., Dinstein I., "Textural Features for Image Classification", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-3, No. 6, 610–621 (1973).

[20] Thiyaneswaran B., Anguraj K., Kumarganesh S., Thangaraj K., "Early detection of melanoma images using gray level co-occurrence matrix features and machine learning techniques for effective clinical diagnosis", *International Journal of Imaging Systems and Technology*, No. 31, 682–694 (2021), DOI: 10.1002/ima.22514.

[21] Sferrazza P., "Grey level co-occurrence matrix and learning algorithms to quantify and classify use-wear on experimental flint tools", *Journal of Archaeological Science: Reports*, Vol. 48, art. 103869 (2023), DOI: 10.1016/j.jasrep.2023.103869.

[22] Hadid A., "The Local Binary Pattern Approach and its Applications to Face Analysis", *2008 First Workshops on Image Processing Theory, Tools and Applications*, p. 1–9, Sousse, Tunisia, 2008.

[23] Hu Y., Long Z., Al Regib G., "Multi-Level Texture Encoding and Representation (Multer) Based on Deep Neural Networks", *2019 IEEE International Conference on Image Processing* (*ICIP*), p. 4410–4414, Taipei, Taiwan, 2019.

[24] Xu X., Li Y., Wu Q.M.J., "A compact multi-pattern encoding descriptor for texture classification", *Digital Signal Processing*, Vol. 114, art. 103081 (2021), DOI: 10.1016/j.dsp.2021.103081.

Przegląd technik wykrywania cech obiektu oraz osadzania ich w przestrzeniach wielowymiarowych

R. WASILUK

Każdy obiekt w otaczającym nas świecie posiada unikalne cechy. Część z nich, np. liczba kół w pojeździe, powierzchnia nieruchomości lub rok budowy budynku, może być z łatwością opisana przy wykorzystaniu liczb. Istnieją jednak takie cechy obiektów, które charakteryzują się wyższą złożonością i nie można ich w tak prosty sposób opisać. Przykładem takich właściwości może być: kształt, zestaw kolorów lub faktura materiału obiektu. Cechy te są definiowane w kontekście obiektów przedstawionych na obrazach i opisują najbardziej podstawowe właściwości obiektów świata rzeczywistego. Tematyka przetwarzania wspomnianych cech wizualnych jest tematem badań naukowych od wielu dekad, a powstała w tym zakresie literatura jest bardzo obszerna. Niniejszy artykuł zbiera najważniejsze z istniejących podejść do zagadnienia wykrywania kształtu obiektu, jego kolorów, faktury materiału oraz osadzania reprezentacji w przestrzeniach wielowymiarowych. Wykorzystanie tych metod czyni możliwym przedstawienie cech obiektu pod postacią punktów w przestrzeniach wielowymiarowych. To z kolei otwiera drogę do użycia przygotowanych reprezentacji w rozwiązywaniu zadań optymalizacji wielokryterialnej.

Słowa kluczowe: cechy wizualne, ekstrakcja cech, osadzanie cech.