Artificial Intelligence Vision Based on Computer Digital Technology in 3D Image Colour Processing

Tianshuang Zhang¹, * and Yunfeng Ma ²

¹Shenyang Jianzhu University, Shenyang, China,
²Shenyang Aerospace University, Shenyang, China

zhangtianshuang@sjzu.edu.cn,

Abstract. Based on computer digital technology, the paper first uses the method of non-equal interval quantization to quantize the pixels in the 3D image, and introduces the artificial intelligence vision technology HSV colour space to construct the feature vector formula after non-equal interval quantization to calculate the difference between the two images the final distance between. Experiments have shown that the algorithm has better robustness to changes in light intensity while maintaining the lower number of quantized bins in the Tahoun algorithm. At the same time, the algorithm can effectively reduce 3D colours, with better visual effects, and better than the original. The interference of images with large differences on the search results improves the recall rate of the image library used in the experiment.

Key words. Computer digital technology, artificial intelligence, machine vision, 3D image colour, RGB, HSV, tahoun algorithm.

1. Introduction

3D target recognition is an important research content in pattern recognition and machine vision. It can be applied in many fields such as military, industry, and biomedicine, and has great practical value and important significance. At present, the main problems of 3D target recognition based on two-dimensional RGB images are: 1) It is easily affected by external conditions such as lighting; 2) During the image projection process, the 3D structure information of the actual scene is lost, resulting in the inability to extract and characterize 3D objects. The characteristics of the essential attributes; 3) It is difficult to solve the 3D target recognition problem of intra-class differences and similarity between classes, and the target recognition accuracy is not high.

At present, the rapid detection methods of visually significant objects are mainly studied from two perspectives. One is based on high-level visual features, tasks, upper-level knowledge, etc., which imitate human top-down visual salient object detection models. This method is based on a detection method based on distinguishing saliency, clustering the extracted pixel block features, and the model matching library is established as a priori knowledge to simulate the human eye's ability to distinguish different objects. Some scholars have proposed a target detection method based on the perceptual domain histogram (RFCH). This method uses RFCH to build the appearance model of the object, and comprehensively uses descriptors such as normalized color, gradient amplitude and Laplacian, and has certain adaptability to scale, illumination and object occlusion [1]. The other is a bottom-up visually
significant object detection model based on low-level visual features and directly using information in
the scene for detection. Some scholars have proposed a spatial domain-based calculation model for
visual saliency objects. The basic idea is to form feature maps on each feature dimension based on visual
features such as grayscale, color, and direction, and then analyze and merge to obtain the final saliency
map. Some scholars have proposed a context-aware (CA) saliency detection method, which considers
both the local and global features of the image. Overcoming the salient area range is the fixed model
and the area only considers the foreground graphics and ignores the background with information
content. The outline of the salient area can be extracted, which is conducive to subsequent processing,
but the saliency of each pixel in the image relative to the local area needs to be calculated. The amount
of calculation is large.

Based on the three primary color depth sensor (RGB-D) to obtain indoor scene images, this paper
proposes a rapid detection method for visually significant objects in indoor three primary colors (RGB)
images that integrates depth information [2]. This method effectively reduces the computational
complexity of the traditional visual attention mechanism for image saliency detection, improves
the detection accuracy, and achieves the purpose of rapid and accurate detection of indoor salient objects.

2. Quantification of 3D image color space

We generally get the RGB value from the image first, so we need to convert the RGB value to HSV
value. Then the three components of H, S, and V are quantized at unequal intervals according to human
color perception. The details are as follows:

According to the human visual resolution ability, the three components of the HSV color space are
respectively quantized, and the quantization levels of the H, S, and V components are 8, 3, and 3,
respectively [3]. Quantify based on the difference in color and subjective color perception.

\[
H = \begin{cases} 
0 & \text{if } h \in [316, 20] \\
1 & \text{if } h \in [21, 40] \\
2 & \text{if } h \in [41, 75] \\
3 & \text{if } h \in [76, 155] \\
4 & \text{if } h \in [156, 190] \\
5 & \text{if } h \in [191, 270] \\
6 & \text{if } h \in [271, 295] \\
7 & \text{if } h \in [296, 315] 
\end{cases}
\]

\[
S = \begin{cases} 
0 & \text{if } s \in [0, 0.2] \\
1 & \text{if } s \in [0.2, 0.7] \\
2 & \text{if } s \in [0.7, 1] 
\end{cases}
\]

\[
V = \begin{cases} 
0 & \text{if } v \in [0, 0.2] \\
1 & \text{if } v \in [0.2, 0.7] \\
2 & \text{if } v \in [0.7, 1] 
\end{cases}
\]

According to optical theory, the color of an object is related to the wavelength or frequency of light.
Different colored lights have different wavelengths and frequency ranges in vacuum, so the tones can
be quantified at unequal intervals. Construct a feature vector [4]. According to the above quantization
level, the 3 color components are synthesized into a feature vector

\[
G = HQ_vQ_s + SQ_v + V
\]

Among them, \(Q_v\) and \(Q_s\) are the quantization stages of the components S and V, here we take \(Q_v = 3,\)
\(Q_s = 3\). Then this formula becomes
In this way, the three components of the HSV colour space are described as a feature vector. According to the value range of the above formula G is \([0,1,...,71]\), calculate G to obtain a one-dimensional histogram of 72 bins. The cumulative histogram vector \(\overline{H}\) of the image can be defined as:

\[
\overline{H} = \{\overline{h}[c_1], \overline{h}[c_2],...,\overline{h}[c_k],...,\overline{h}[c_n]\}, 0 \leq \overline{h}[c_k] \leq 1
\]

\(\overline{h}[c_k]\) represents the cumulative frequency of pixels of \(c_i - c_k\) colors:

\[
\overline{h}[c_k] = \sum_{c_i \leq c_k} h[c_i]
\]

The cumulative color histogram can be calculated using the color histogram:

\[
\overline{H} = \{\overline{h}[0], \overline{h}[1],...,\overline{h}[c],...,\overline{h}[n]\}
\]

According to the definition of the cumulative color histogram, \(\overline{h}[n]\) is always equal to 1, so it is meaningless to calculate the similarity distance, so the value of \(c\) is 0, 1, 2, ..., n-1, so the final cumulative histogram feature vector is:

\[
\overline{H} = \{\overline{h}[0], \overline{h}[1],...,\overline{h}[c],...,\overline{h}[n-1]\}
\]

Experiments have proved that because the zero value in the histogram is eliminated, the effect of the cumulative colour histogram intersection is better than the simple histogram intersection.

3. 3D visual colour saliency detection

Salient object detection based on spectral residual and inverse Fourier transform-The model mainly obtains salient attention areas by detecting areas with sudden changes in intensity in the image and eliminating redundant background information [5]. The model defines the logarithmic spectrum \(L(f)\) of the image in the frequency domain as the overall information \(H(I_n)\) of the image. Use the average template \(H_a(f)\) to smooth the logarithmic spectrum \(L(f)\) to obtain the average spectrum \(M[L(f)]\), namely

\[
M[L(f)] = H_a(f) * L(f)
\]

In the formula, \(H_a(f)\) is the mean filter template of nun, and * represents the convolution operation. Subtract \(S(f) = L(f) - M[L(f)]\) from the logarithmic spectrum and the average spectrum to obtain the spectral residual \(S(f)\). According to equation (10), perform the inverse Fourier transform on the sum of the spectral residual \(S(f)\) and the phase spectrum \(P(f)\) to obtain the final saliency map \(S(x)\), namely

\[
S(x) = g(x) * F_{irre} \left\{ \exp \left[ S(f) + P(f) \right] \right\}^2
\]

Where \(g(x)\) is a Gaussian filter. Based on the above research theory, we propose a multi-feature visual attention mechanism saliency detection model. The calculation process is shown in Figure 1. This model mainly detects areas with large colour contrast and large brightness changes in the image, and obtains a saliency map of the visual attention mechanism.
Figure 1. Saliency map acquisition flowchart

4. Design of 3D colour processing system based on artificial vision

4.1. System design

The on-line selection and sorting system based on object colour characteristics designed in this paper is mainly composed of two major subsystems: machine vision pattern recognition system and mechanical sorting control system. An image collector (industrial camera) and a computer are used to form a host computer system for image classification based on object colour. The lower computer system is composed of feeding device, transmission device, sorting device and PLC control device. After the upper computer performs image acquisition and processing and pattern recognition, the recognition result is transmitted to the lower computer [6]. According to the recognition result, the lower computer controls the three cooperating mechanical devices through the PLC control device to complete the product loading, transmission, sorting and warehousing work, to realize the automatic operation of product sorting. The design architecture is shown in Figure 2.

Figure 2. 3D color recognition artificial intelligence system
Generally, the content-based image classification process includes digital image reading, digital image processing, feature extraction, image classification and result output. The host computer uses an industrial camera equipped with a USB interface and an Ethernet interface to take pictures and samples of the object samples, and transfer the collected images to a designated path in the computer for digital image reading. The literature shows that MATLAB 7.0 software not only performs well in graphics processing functions and general data visualization software functions (such as the drawing and processing of two-dimensional curves and three-dimensional curved surfaces, etc.), but also for functions that other software does not have (such as graphics lighting). Processing, chromaticity processing and performance of four-dimensional data, etc.) also have excellent processing capabilities. Therefore, pattern recognition algorithms such as digital image processing, feature extraction, image classification and result output are implemented using the computer system software MATLAB.

4.2. Colour space extraction

Before we extract the color of the object, we first need to establish a color space, and use a mathematical method-color space to represent the color, and use it to accurately describe the color and visualize the color. According to different application objects, the color space in digital image engineering mainly includes the color space RGB and CMY color space for hardware devices; the color space HSI for color image processing; the CIE color space, L*a*b* and XYZ color space. The color image obtained by the industrial camera is directly expressed as the RGB color space, but the red (R), green (G), and blue (B) components are not directly related to human perception of color, while L*a*b* color in the three-dimensional space, the visual color and the psychological perception color are equidistant on each coordinate axis, and there is a fine level difference between the lightness index and the chromaticity index, and the color difference resolution is very strong [7]. It is especially suitable for color measurement and measurement in the case of small color difference. Distinguish. Therefore, this design uses L*a*b* color space for feature value extraction. Using the custom function 'RGB2Lab' in MATLAB, the RGB color space can be easily converted to the L*a*b* color space. As described in section 2.1 above.

5. Experimental Design

In order to further examine the performance of the algorithm, the images in the image library are taken as example images in turn, and the most similar 30 images are returned for each retrieval, and the algorithm's average recall rate, average precision rate, and highest search rate for each category of images are calculated. The full rate and minimum recall rate are shown in Table 1 and Table 2 below

| Table 1. Search results of Tahoun algorithm |
|--------------------------------------------|
| Average recall rate | Highest recall rate | Minimum recall rate |
|---------------------|---------------------|---------------------|
| Aircraft            | 41.0%               | 45.0%               | 10.0%               |
| Yellow flower       | 82.0%               | 100.0%              | 70.0%               |
| Human face          | 74.5%               | 95.0%               | 10.0%               |
| Horse               | 55.3%               | 70.0%               | 30.0%               |
| Coastal             | 45.8%               | 65.0%               | 20.0%               |
| Board               | 55.3%               | 70.0%               | 20.0%               |

| Table 2. Retrieval effect of HSV spatial component distance comprehensive algorithm |
|-------------------------------------------------------------------------------------|
| Average recall rate | Highest recall rate | Minimum recall rate |
|---------------------|---------------------|---------------------|
| Aircraft            | 62.3%               | 75.0%               | 30.0%               |
| Yellow flower       | 99.8%               | 100.0%              | 95.0%               |
| Human face          | 75.8%               | 95.0%               | 30.0%               |
| Horse               | 37.3%               | 60.0%               | 20.0%               |
| Coastal             | 63.3%               | 75.0%               | 5.0%                |
| Board               | 84.5%               | 95.0%               | 20.0%               |
It can be seen from the experimental data that for 5 of the 6 categories, the HSV spatial component distance synthesis algorithm proposed in this paper has a significant improvement in the average recall rate, the highest recall rate, and the lowest recall rate. The category has been reduced. It can be seen that the HSV spatial component distance synthesis algorithm proposed in this paper is better than Tahoun algorithm in terms of overall retrieval performance.

6. Conclusion

The paper proposes a rapid detection method for visually significant objects in RGB images under the fusion of computer artificial intelligence and machine vision. Compared with the traditional detection method, the saliency map of the RGB image is obtained through the multi-feature visual attention mechanism saliency detection model, which not only reduces the computational complexity of the attention mechanism, but also improves the detection efficiency. In addition, the use of the salient region growth strategy solves the shortcomings of inaccurate and missing segmentation of traditional salient regions. Through the fusion of depth information, the number of objects in the salient area and the positional relationship between each other are obtained.

References

[1] Azarafza, M., Ghazifard, A., Akgün, H., & Asghari-Kaljahi, E. Development of a 2D and 3D computational algorithm for discontinuity structural geometry identification by artificial intelligence based on image processing techniques. Bulletin of Engineering Geology and the Environment, 78(5) (2019) 3371-3383.

[2] Bayoudh, K., Hamdaoui, F., & Mtibaa, A. Transfer learning based hybrid 2D-3D CNN for traffic sign recognition and semantic road detection applied in advanced driver assistance systems. Applied Intelligence, 51(1) (2021) 124-142.

[3] Chen, D., Li, G., Sun, Y., Jiang, G., Kong, J., Li, J., & Liu, H. Fusion hand gesture segmentation and extraction based on CMOS sensor and 3D sensor. International Journal of Wireless and Mobile Computing, 12(3) (2017) 305-312.

[4] Meka, A., Haene, C., Pandey, R., Zollhöfer, M., Fanello, S., Fyffe, G., ... & Rheumann, C. Deep reflectance fields: high-quality facial reflectance field inference from colour gradient illumination. ACM Transactions on Graphics (TOG), 38(4) (2019) 1-12.

[5] Liu, Y., Han, F., Li, F., Zhao, Y., Chen, M., Xu, Z., ... & Qian, L. Inkjet-printed unclonable quantum dot fluorescent anti-counterfeiting labels with artificial intelligence authentication. Nature communications, 10(1) (2019) 1-9.

[6] Benke, K., & Benke, G. Artificial intelligence and big data in public health. International journal of environmental research and public health, 15(12) (2018) 2796-2809.

[7] Nakashima, H., Kawahira, H., Kawachi, H., & Sakaki, N. Artificial intelligence diagnosis of Helicobacter pylori infection using blue laser imaging-bright and linked colour imaging: a single-center prospective study. Annals of gastroenterolog