

Research Article

Accurate Recognition of Motion Patterns Based on Artificial Visual Neural Network

Meiqi Li 

Hubei University of Technology, Wuhan 430070, China

Correspondence should be addressed to Meiqi Li; 2010221129@hbut.edu.cn

Received 15 April 2022; Accepted 17 May 2022; Published 4 July 2022

Academic Editor: Qiangyi Li

Copyright © 2022 Meiqi Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the accuracy of motion pattern recognition, this paper combines the artificial visual neural network to construct a motion pattern recognition system. Moreover, this paper discusses the psychological perception properties of human eyes to color stimuli and gives a description of the observation field of view where the color stimuli are located. At the same time, this paper analyzes the phenomenon of color adaptation and provides a method of color appearance matching through modeling to achieve color appearance matching under variable observation conditions. Based on the chromatic adaptation transformation, a chromatic appearance model is given, which can predict the corresponding color and also predict the chromatic appearance properties of color stimuli under given observation conditions. In addition, this paper constructs an intelligent motion pattern recognition system combined with artificial visual neural network. The experimental results show that the motion pattern recognition system based on artificial visual neural network can accurately identify the motion pattern category.

1. Introduction

Trajectory resampling refers to making the time interval between consecutive sampling points in the trajectory the same by inserting the mean or median value among the sampling points. Trajectory segmentation refers to segmenting the trajectory at important turning points by MDL value or DP algorithm and dividing a trajectory into multiple trajectory segments for analysis and mining. Trajectory feature extraction refers to extracting features with strong identification ability from the trajectory. The quality of feature extraction largely determines the accuracy of motion pattern recognition and is a key step in trajectory motion pattern recognition. In addition, the extracted trajectory features generally include trajectory motion features, shape features, location features where the trajectory appears, and time features. The classifier refers to constructing a classifier based on the trajectory data of the known category. The model can predict the trajectory data of the unknown category and obtain one of the given categories. Establishing a classifier means using the feature vector of the training trajectory as the input of the classifier to train a classifier for

motion pattern recognition. In the test phase, the feature vector of the test trajectory is used as input, and the motion pattern category of the test trajectory is obtained through the trained classifier. Currently, commonly used classifiers include k-nearest neighbors, decision trees, random forests, neural networks, support vector machines, Bayesian networks, conditional random fields, etc.

With the popularization of electronic devices with global positioning system functions, the behavior characteristics and laws of individuals or groups can be analyzed by collecting the location information of moving objects. It can provide an important basis for commanding decision-makers in the fields of processing and large-scale military operations. Group movement pattern analysis is to analyze the movement state of the group and extract its behavior pattern through the information of the position of each member in the group changing over time, so as to comprehensively evaluate the group behavior pattern and action effect. The process of group motion pattern analysis mainly includes the steps of data preprocessing, constructing time series, using interpolation algorithm to calculate the corresponding member positions of the time series, and

calculating various description parameters for data analysis. Among them, data preprocessing mainly includes coordinate transformation, time synchronization, and exclusion of abnormal nodes. Since each member node records and reports its position information at its own rhythm, the original data set generally does not contain the position data of all nodes at a certain time during group motion pattern analysis. It is necessary to construct a unified time series of the group on the basis of the time series of each member and then use the interpolation algorithm to determine the position information of the member nodes in the group at this moment under the time series.

In this paper, the artificial visual neural network is combined to construct a motion pattern recognition system to improve the effect of motion pattern recognition and effectively improve the accuracy of motion pattern recognition.

2. Related Work

Human body motion pattern recognition refers to the process of identifying various motion states of the human body, mainly by analyzing the data of inertial sensors to identify motion patterns. Motion pattern recognition technology is widely used in competitive sports, health detection, medical research and pedestrian navigation, and other fields [1]. In recent years, it has also been used in some rescue work [2]. Literature [3] placed accelerometers on the waist and thighs and used wavelet transform to analyze postures such as walking, jogging, and lying down. The experimental results were compared with video recordings, indicating that the data from the accelerometer can be used to identify human posture. Literature [4] proposes a recognition algorithm in which the accelerometer is placed on the wrist and compares and analyzes the recognition results of human behavior patterns on the wrist and buttocks. Literature [5] places the sensor motion node on the front right hip and proposes a new human action recognition framework based on compressed sensing and sparse representation theory. Literature [6] uses decision tree and logistic regression technology to propose a new prediction model based on machine classification learning. The recognition results of running and sitting are good, but the recognition results of going up and down the stairs are not very good. Literature [7] analyzed the data of two inertial measurement units placed on the foot and shoulder to realize indoor pedestrian navigation. Literature [8] uses multiple sensors placed at different positions of the body and proposes a classification and recognition algorithm suitable for wearable sensor platforms, whose accuracy is higher than that of a single sensor.

It is proposed that current anomaly detection methods can be mainly divided into two categories: trajectory-based methods [9] and appearance feature-based methods [10]. The former, as the most traditional anomaly detection method, usually consists of two steps, tracking the target in the video scene to obtain the motion trajectory and then modeling and analyzing the tracking trajectory. In complex scenes with many targets, both target tracking and trajectory analysis are

difficult, resulting in high computational cost. The most common technique to solve this problem is to use the method of feature extraction, which proposes a visual feature called optical flow texture and combines it with spatial information to detect abnormal behavior, but this method is only suitable for detection. Behavior differs from normal movement patterns. Optical flow mainly focuses on the motion information of objects and often ignores some nondynamic abnormal information [11]. Some more complete representations are proposed to ensure the inclusion of dynamic and static information. Literature [12] uses optical flow and contour features to generate spatiotemporal descriptions and uses Nonnegative Locality-Constrained Linear Coding (NLLC) to detect abnormal behaviors. Many different kinds of abnormal behavior detection models have emerged in recent years. Literature [13] proposes to segment the context of a video into various semantic regions, establish a semantic context model for moving objects in the context, and then use the model to detect abnormal behavior. Literature [14] proposed a joint sparse model (Jointly Sparse Model, JSM) to train the trajectory in the training sample to obtain an overcomplete dictionary and use the dictionary to sparsely reconstruct the trajectory of the test sample and then detect anomalies through the reconstruction error. Topic models can identify behavior patterns in scenes and detect anomalies through underlying features that appear at different levels and have achieved great success in the field of behavior recognition. Literature [15] utilizes hierarchical Bayesian models, such as Latent Dirichlet Allocation (LDA) and Hierarchical Dirichlet Process (HDP), to describe typical behaviors in videos. Literature [16] proposed a trajectory analysis method based on the LDA model and used it for abnormal behavior detection. Literature [17] proposed to use the Probabilistic Latent Semantic Analysis (PLSA) method to build a topic model with local information and quantify the location and size information of the image through rich spatiotemporal gradient descriptors to extend topic-based analysis and local descriptors, supplemented with some easy-to-ignore exception local information. However, traditional probabilistic topic models lack a mechanism to directly control the sparsity of document representation. Literature [18] proposed a sparse local coding (Sparse Topical Coding, STC) method to find latent representations in a large dataset and directly control the sparsity of model representations through sparsity-inducing regularization terms. A method for detecting abnormal video behavior based on motion pattern analysis is proposed. In the underlying processing process, the spatiotemporal descriptors are extracted and combined with position information to generate visual words, so that the visual words contain sufficient dynamic and static information, which can be used for the detection of various abnormal behaviors [19].

Previous related research work mainly focused on extracting the motion features of trajectories, including velocity, acceleration, and motion direction. If only the motion features of the trajectory are extracted, the important location features of the trajectory are often ignored. In some special cases, such as in traffic jams, the motion characteristics of multiple motion modes are similar and difficult to distinguish. Literature [20] believes that the location

information of the trajectory can help researchers improve the accuracy of trajectory motion pattern recognition, so it proposes a region classification rule considering duration and a path classification rule mining algorithm considering duration.

Due to the large drift and low accuracy of MEMS devices, the auxiliary means of Zero Velocity Update (ZUPT) are usually used to suppress the accumulation of errors, and the solution accuracy is further improved by periodic error clearing. In the process of walking, the time when the feet remain relatively stationary with the ground is very short. Therefore, the detection of the zero-speed interval is not only the core of the zero-speed correction, but also the key link of the pedestrian navigation solution [21]. There may be missed judgments or misjudgments during zero-speed zone detection. Missing judgment means that all zero-speed state points are not judged, and misjudgment means that non-zero-speed state points are judged to be zero-speed state. Since the boundary point between zero-speed and non-zero-speed states cannot be accurately determined, the occurrence of missed judgments is inevitable [22].

3. Machine Vision Feature Recognition Algorithms

If the apparent brightness of a color stimulus is B , the brightness is L , and the apparent brightness of the reference white point is ∞ , the relationship between them can be expressed as

$$L = \frac{B}{B_w}. \quad (1)$$

If the chroma of a color stimulus is M , the chroma is C , and the chroma of the reference white point (same as the brightness) is B_w , the relationship between them can be expressed as

$$C = \frac{M}{B_w}. \quad (2)$$

Saturation indicates the purity of the color or the difference from a neutral. If hue is the perception of the dominant wavelength, saturation is the degree to which other wavelengths of light are doped in the dominant wavelength. The wider the range of wavelengths contained in the color is, the less saturated the resulting color will feel.

If the saturation of a color stimulus is recorded as S , then

$$S = \frac{M}{B} = \frac{C}{L}. \quad (3)$$

The observation condition attribute is used to describe the scene (Scene) in which a color stimulus is observed. The scene is often referred to as the viewing field or viewing conditions. Viewing conditions have a large impact on color perception. This subsection will define a simple typical observation field (Figure 1), which consists of four parts: the color stimulus (Stimulus), the proximal field (Proximal field), the background (Background), and the surrounding environment (Surround).

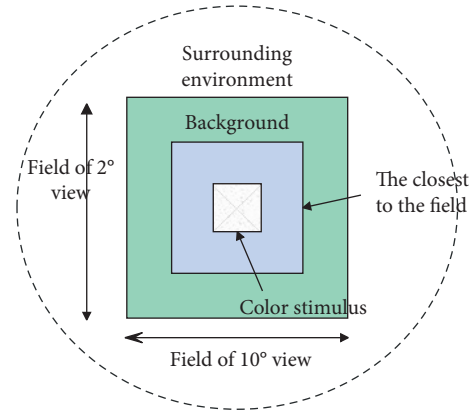


FIGURE 1: A typical observation field.

The phenomenon of spatial structure refers to the phenomenon that color appearance changes with the spatial structure and background of color stimuli. The most famous phenomenon is simultaneous contrast, that is, the phenomenon that color appearance changes to its opposite direction under a contrast perception. If a brighter background induces a stimulus to appear darker, a darker background will induce a brighter appearance. In the same way, red induction produces green, green induction produces red, yellow induction produces blue, and blue induction produces yellow. Figure 2 shows how the color appearance changes when grays of the same shade are placed against different backgrounds. Among them, the two color blocks above are placed on the same gray background, and the visual perception is the same. However, when they are placed against the different backgrounds below, the grays placed against the black background are visually perceived as “brighter,” while the grays placed against the white background feel slightly darker. Figure 3 further summarizes the strong effect of changes in the background on color appearance. In the figure, the background is getting brighter and brighter from left to right, and the chromaticity values on each color bar are exactly the same, but in visual perception. The color difference from left to right is felt due to the gradient of the background color.

The change in the sensitivity of the human eye caused by the change of the illumination light can be shown in Figure 4, which reflects the situation of changing from sunlight illumination to incandescent lamp illumination. In daylight illumination, since the spectral distribution of sunlight is roughly flat, the red, green, and blue sensitivities are roughly balanced. When changing to incandescent lighting, the red component increases and the blue component decreases, so the sensitivity of the red photoreceptor decreases and the blue photoreceptor increases, resulting in always a fixed response. This explains why the color appearance does not change.

According to the von Kries hypothesis, its color adaptation model can be established through the following steps.

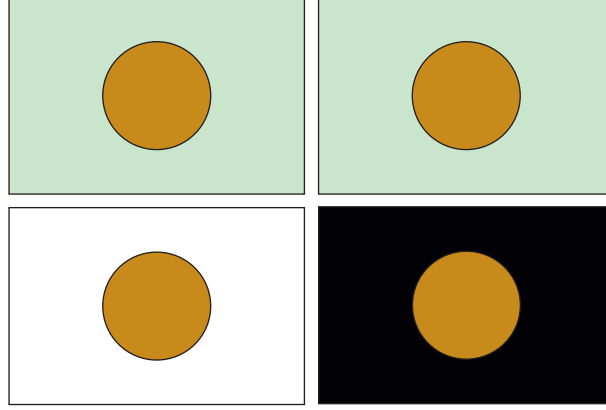


FIGURE 2: Simultaneous comparison phenomenon 1.

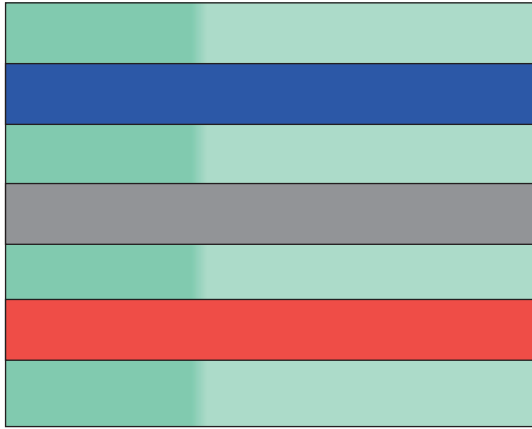


FIGURE 3: Simultaneous comparison phenomenon 2.

- (1) First, the algorithm converts the CIEXYZ tristimulus value to the LMS cone response value (sometimes also called RGB or response), and the conversion relationship is

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = M \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (4)$$

$$M = \begin{pmatrix} 0.4002 & 0.7076 & -0.0808 \\ -0.2263 & 1.1653 & 0.0457 \\ 0.0 & 0.0 & 0.9182 \end{pmatrix}.$$

Among them, M represents the Hunt-Pointer-Estevéz transformation matrix, and its elements are determined by the chromaticity coordinates of the three basic primary colors.

- (2) By multiplying the initial vertebral response LMS of the human eye to color stimuli and the independent gain control coefficient $\alpha_L, \alpha_M, \alpha_S$, the adaptive vertebral signal LMS_{adapted} is obtained. The gain control coefficient is the most critical part of most color adaptation models. In the von Kries model, the reciprocal of the

scene's maximum LMS_{max}-pheasant response value is used as the independent gain control coefficient. Usually, the maximum LMS pheasant response of the scene is the pheasant response LMS_{white} of the scene white point, so the von Kries adaptation is usually also called white point adaptation and has [23]

$$\begin{bmatrix} L_{\text{adapted}} \\ M_{\text{adapted}} \\ S_{\text{adapted}} \end{bmatrix} = \begin{bmatrix} \alpha_L & 0.0 & 0.0 \\ 0.0 & \alpha_M & 0.0 \\ 0.0 & 0.0 & \alpha_S \end{bmatrix} \cdot \begin{bmatrix} L \\ M \\ S \end{bmatrix}. \quad (5)$$

Among them,

$$\alpha_L = \frac{1}{L_{\text{max}}} \text{ or } \alpha_L = \frac{1}{L_{\text{white}}},$$

$$\alpha_M = \frac{1}{M_{\text{max}}} \text{ or } \alpha_M = \frac{1}{M_{\text{white}}}, \quad (6)$$

$$\alpha_S = \frac{1}{S_{\text{max}}} \text{ or } \alpha_S = \frac{1}{S_{\text{white}}}.$$

- (3) Through the adapted pheasant signal LMS_{adapted}, the algorithm calculates the adapted color stimulus value XYZ_{adpoted} as follows:

$$\begin{bmatrix} X_{\text{adpoted}} \\ Y_{\text{adpoted}} \\ Z_{\text{adpoted}} \end{bmatrix} = M^{-1} \cdot \begin{bmatrix} L_{\text{adpoted}} \\ M_{\text{adpoted}} \\ S_{\text{adpoted}} \end{bmatrix}. \quad (7)$$

For any given color stimulus, it is very useful to obtain the adapted CIEXYZ value, but more often it is more important to obtain the corresponding color of the color stimulus under another observation condition. Chromatic Adaptation Transform (CAT) is used to calculate the corresponding color. The derivation of the von Kries chromatic adaptation transformation is as follows.

The color stimulus value under the source observation condition is XYZ, and the corresponding color is X₂Y₂Z₂ under the target observation condition. The task of von Kries color adaptation transformation is to match the color

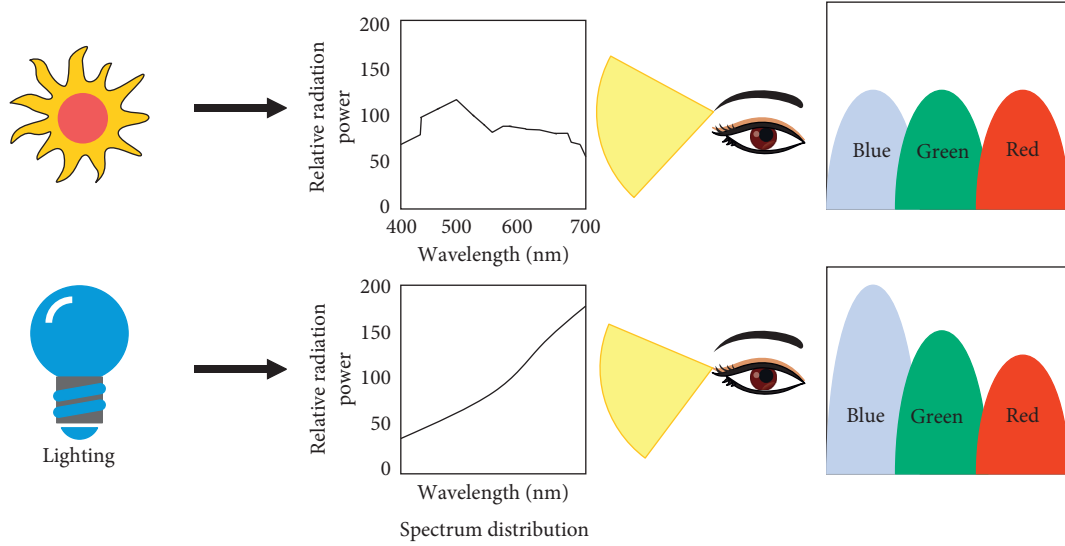


FIGURE 4: Changes in the balance of human eye sensitivity due to changes in illumination light.

appearance of two color stimuli under the source and destination observation conditions, that is, to make the adaptive vertebral response signals equal under the two observation conditions.

For the convenience of expression, we represent this transformation as a linear matrix. From the von Kries model introduced above, we can see that the adaptive response value LMS_{adapped} of the pheasant under the source observation condition is

$$\begin{bmatrix} L_{\text{adpsted}} \\ M_{\text{adapped}} \\ S_{\text{adpped}} \end{bmatrix} = \begin{bmatrix} \frac{1}{L_{\text{white}}} & 0.0 & 0.0 \\ 0.0 & \frac{1}{M_{\text{white}}} & 0.0 \\ 0.0 & 0.0 & \frac{1}{S_{\text{white}}} \end{bmatrix} \cdot M \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (8)$$

The response value $L_2M_2S_2$ of the pheasant after adaptation under the objective observation condition is

$$\begin{bmatrix} L_{2\text{adapted}} \\ M_{2\text{adapted}} \\ S_{2\text{adapted}} \end{bmatrix} = \begin{bmatrix} \frac{1}{L_{\text{white}2}} & 0.0 & 0.0 \\ 0.0 & \frac{1}{M_{\text{white}2}} & 0.0 \\ 0.0 & 0.0 & \frac{1}{S_{\text{white}2}} \end{bmatrix} \cdot M \cdot \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix}. \quad (9)$$

Among them, $L_{\text{white}2}$, $M_{\text{white}2}$, and $S_{\text{white}2}$ represent the cone response value of the target observation white point, so the corresponding color $X_2Y_2Z_2$ under the target observation condition is

$$\begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = M^{-1} \cdot \begin{bmatrix} \frac{L_{\text{white}2}}{L_{\text{white}}} & 0.0 & 0.0 \\ 0.0 & \frac{M_{\text{white}2}}{M_{\text{white}}} & 0.0 \\ 0.0 & 0.0 & \frac{S_{\text{white}2}}{S_{\text{white}}} \end{bmatrix} \cdot M \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (10)$$

This paper proposes a model that combines the linear process established by the von Kries model and the non-linear process established by the exponential adaptation, which is called the Nayatani model.

When calculating the adaptive cone response signal LMS_{adapted} , the Nayatani model uses a nonlinear model to multiply the gain adjustment coefficient by an episodic function. The exponent of the curtain function is based on a variable that adapts to the field brightness. This enables the Nayatani model to reflect luminance phenomena such as Hunt and Stevens effects. In addition, a noise control function is added for threshold prediction. The gain adjustment factor is used to control the brightness of the nonselective samples (gray) as the adaptation field brightness to avoid complete color constancy. The specific formula is as follows:

$$\begin{aligned} L_{\text{adpsted}} &= \alpha_L \cdot \left(\frac{L + L_n}{L_{\text{white}} + L_n} \right)^{\beta_L}, \\ M_{\text{adpsted}} &= \alpha_M \cdot \left(\frac{M + M_n}{M_{\text{white}} + M_n} \right)^{\beta_M}, \\ S_{\text{adpsted}} &= \alpha_S \cdot \left(\frac{S + S_n}{S_{\text{white}} + S_n} \right)^{\beta_S}. \end{aligned} \quad (11)$$

Among them, L_n , M_n , and S_n represent the added additive noise, β_L , β_M , and β_S are variables based on adaptive brightness level, and α_L , α_M , and α_S are the coefficients that make neutral gray stimuli produce color constancy.

The basic structure of the Fairchild model is consistent with the von Kries model, and only the method of calculating the gain control coefficient is more complicated. The formula for calculating the adaptive pheasant response value LMS_{adapted} is as follows:

$$\begin{bmatrix} L_{\text{adapted}} \\ M_{\text{adapted}} \\ S_{\text{adapted}} \end{bmatrix} = \begin{bmatrix} \alpha_L & 0.0 & 0.0 \\ 0.0 & \alpha_M & 0.0 \\ 0.0 & 0.0 & \alpha_S \end{bmatrix} \cdot \begin{bmatrix} L \\ M \\ S \end{bmatrix}. \quad (12)$$

The formula for calculating coefficient α_L is as follows (the other two coefficients α_M and α_S are calculated in a similar way):

$$\alpha_L = \frac{p_L}{L_n},$$

$$p_L = \frac{1 + Y_n^{1/3} + l_E}{1 + Y_n^{1/3} + 1/l_E}, \quad (13)$$

$$l_E = \frac{3(L_n/L_E)}{L_n/L_E + M_n/M_E + S_n/S_E}.$$

Among them, Y_n refers to the brightness of the adaptation field, and the unit is cd/m^2 , L_n , M_n , and S_n represent the response value of the pheasant adapted to the white spot stimulus, and L_E , M_E , and S_E represent the response value of the pheasant with equal energy illumination. When p_L is equal to 1 (the adaptive white point is equal energy illumination $L_n = M_n = S_n$, that is, $l_E = 1$), it is considered to be fully adapted, and the model is completely transformed into the von Kries model. When p_L is L_n and α_L is 1, it is considered to be zero adaptation. When the p_L value is between the two, it belongs to incomplete adaptation. The degree of adaptation depends on the degree of adaptation to the field brightness level and the degree of adaptation to energy illumination such as white point deviation. This color-adaptive conversion method is similar to the von Kries conversion method described above.

The main difference between the spectral sharpening chromatic adaptation model and the abovementioned chromatic adaptation model lies in the first step of chromatic adaptation, that is, the calculation method of the pheasant response value. This type of model first normalizes the CIEXYZ tristimulus values by dividing by the component. The obtained cone response value RGB does not represent the physiological pheasant response signal, but the spectral sharpened pheasant response signal, which can often better maintain color saturation and color constancy. A specific spectral sharpening model, the Bradford (BFD) chromatic adaptation model, implements the following steps.

- (1) The algorithm normalizes the CIEXYZ tristimulus values and converts them to the RGB spectral

sharpening cone response space through the BFD matrix. The formula is

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_B \begin{bmatrix} \frac{X}{Y} \\ \frac{Y}{Y} \\ \frac{Z}{Y} \end{bmatrix}, \quad (14)$$

$$M_B = \begin{pmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{pmatrix}.$$

Among them, the elements in the M_B matrix are obtained by optimizing the Lam&Rigg dataset.

- (2) The algorithm calculates the adaptive spectral sharpening pheasant response value $R_c G_c B_c$.

The model uses a nonlinear power function for short waves, and the calculation formula is as follows:

$$R_c = \left[D \left(\frac{1.0}{R_w} \right) + 1 - D \right] R,$$

$$G_c = \left[D \left(\frac{1.0}{G_w} \right) + 1 - D \right] G,$$

$$B_c = \left[D \left(\frac{1.0}{B_w^p} \right) + 1 - D \right] |B|^p, \quad (15)$$

$$p = \left(\frac{B_w}{1.0} \right)^{0.0834}.$$

Among them, R_w , G_w , and B_w represent the spectral sharpening pheasant response values adapted to the white point stimulus, and D represents the degree of adaptation. Usually, D is taken as 1.0 in hard copy, which means full color adaptation; that is, the effect of lighting is completely canceled. In soft copy, D takes 0, which means zero adaptation, and the influence of light is not offset. When it is in a darker condition like projection, D takes an intermediate value, which means that the effect of illumination is partially offset.

The model calculates the tristimulus values after adaptation and the method of color adaptation transformation is similar to the aforementioned von Kries transformation method.

The input parameters of the CIELAB model include two sets of CIEXYZ tristimulus values, one is the color stimulus itself, and the other is the color stimulus value of the reference white point. The model calculation formula is as follows:

$$\begin{aligned}
L &= 116f\left(\frac{Y}{Y_n}\right) - 16, \\
a &= 500\left(f\left(\frac{X}{Y_n}\right) - f\left(\frac{Y}{Y_n}\right)\right), \\
b &= 200\left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right).
\end{aligned} \tag{16}$$

In the formula, the function $f(x)$ is defined as follows:

$$f(x) = \begin{cases} x^{1/3}, & x > 0.008856, \\ 7.787x + \frac{16}{116}, & x \leq 0.008856. \end{cases} \tag{17}$$

Among them, X , Y , and Z are the tristimulus values of color stimuli, $X_n\Delta Y_n\Delta Z_n$ is the tristimulus value of the reference white point, and the normalization makes $Y_n = 100$.

The definition of the color appearance model requires the calculation of at least three properties of lightness, chroma, and hue. The model provides the calculation of lightness L , chroma C_{ab} , and hue h_{ab} , and the a and b coordinate values are calculated using the following formulas:

$$\begin{aligned}
C_{ab} &= \sqrt{a^2 + b^2}, \\
h_{ab} &= \tan^{-1}\left(\frac{b}{a}\right).
\end{aligned} \tag{18}$$

Although the CIELAB model is the prototype of the colored appearance model, it has the following defects:

- (1) Wrong von Kries chromatic adaptation leads to inaccurate chromatic appearance prediction.
- (2) The CIELAB model does not consider most of the color appearance phenomena, such as brightness, surrounding environment, background, offset lighting phenomenon, etc.
- (3) The CIELAB model always assumes 100% white point adaptation and does not consider the phenomenon of incomplete color adaptation.
- (4) The CIELAB model is designed to predict small color differences between similar objects under fixed viewing conditions, so it is difficult to generalize as a color appearance model.
- (5) The hue constancy of the CIELAB model, especially in the blue region, is very important for some image processing techniques such as gamut matching.

The RLAB model first converts the color tristimulus values to the corresponding colors under the reference standard environment (D65, 2° observer, illuminance of 318 cd/m², hard copy media) through color adaptation conversion. Then, it calculates the color appearance properties using the corresponding color. For the corresponding color prediction under other conditions, it can be calculated by the inverse model of this model.

- (1) The algorithm calculates the LMS vertebral response value:

$$\begin{aligned}
\begin{bmatrix} L \\ M \\ S \end{bmatrix} &= M \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \\
M &= \begin{pmatrix} 0.3897 & 0.6890 & -0.0787 \\ -0.2298 & 1.1834 & 0.0464 \\ 0.0 & 0.0 & 1.0000 \end{pmatrix}.
\end{aligned} \tag{19}$$

Among them, M still represents the Hunt-Pointer-Estevéz transformation matrix. The matrix is normalized so that equal vertebral response values ($L=M=S=100$) can be obtained when the color stimuli are equal energy illumination (i.e., $X=Y=Z=100$). White dot tristimulus values $X_wY_wZ_w$ are also converted to $L_wM_wS_w$ pheasant response values.

- (2) The algorithm calculates the pheasant response value LMS_{adpoted} after color adaptation.

Using the color adaptation matrix A , the adaptive pheasant response value LMS_{adapted} is obtained as follows:

$$\begin{aligned}
\begin{bmatrix} L_{\text{adapted}} \\ M_{\text{adapted}} \\ S_{\text{adapted}} \end{bmatrix} &= A \bullet \begin{bmatrix} L \\ M \\ S \end{bmatrix}, \\
A &= \begin{bmatrix} \alpha_L & 0.0 & 0.0 \\ 0.0 & \alpha_M & 0.0 \\ 0.0 & 0.0 & \alpha_S \end{bmatrix}.
\end{aligned} \tag{20}$$

Among them,

$$\begin{aligned}
\alpha_L &= \frac{p_L + D(1.0 - p_L)}{L_W}, \\
p_L &= \frac{1.0 + Y_n^{1/3} + 1_E}{1.0 + Y_n^{1/3} + 1.0/1_E}, \\
1_E &= \frac{3L_W}{L_W + M_W + S_W}.
\end{aligned} \tag{21}$$

The D factor describes the degree to which light is offset in chromatic adaptation. In general, $D=1.0$ in hard copy conditions, indicating complete chromatic adaptation. $D=0$ in soft copy, which means no offset light. When in a darker condition such as projection, D takes an intermediate value, which means that the illumination is partially offset. Its value depends on the specific observation conditions and generally takes 0.5 empirically. $\alpha_M\Delta\alpha_S\Delta p_M\Delta p_S\Delta m_E$ and s_E are calculated similarly.

- (3) The algorithm calculates the corresponding color $X_{\text{ref}}Y_{\text{ref}}Z_{\text{ref}}$ under the reference standard conditions.

The reference standard condition is D65 illumination, 2° observers, and hard copy of the adaptation field illumination 318cd/m²; the calculation formula is as follows:

$$\begin{bmatrix} X_{\text{ref}} \\ Y_{\text{ref}} \\ Z_{\text{ref}} \end{bmatrix} = M^{-1} \cdot A_{\text{ref}}^{-1} \cdot A \cdot M \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (22)$$

$$A_{\text{ref}}^{-1} = \begin{bmatrix} L_{D65} & 0.0 & 0.0 \\ 0.0 & M_{D65} & 0.0 \\ 0.0 & 0.0 & S_{D65} \end{bmatrix}.$$

Among them, $L_{D65}M_{D65}S_{D65}$ represents the pheasant response value of the D65 illumination white point. Therefore, if $R = M^{-1} \cdot A_{\text{ref}}^{-1}$, it is a constant matrix, and

$$R = \begin{pmatrix} 1.9569 & -1.1882 & 0.2313 \\ 0.3612 & 0.6388 & 0.0 \\ 0.0 & 0.0 & 1.0000 \end{pmatrix}. \quad (23)$$

- (4) The algorithm calculates the color appearance attribute.

The related color appearance attribute parameters and calculation methods are as follows:

$$\begin{aligned} L^R &= 100(Y_{\text{ref}})^\sigma, \\ a^R &= 430[(X_{\text{ref}})^\sigma - (Y_{\text{ref}})^\sigma], \\ b^R &= 170[(Y_{\text{ref}})^\sigma - (Z_{\text{ref}})^\sigma], \\ c^R &= \sqrt{(a^R)^2 + (b^R)^2}, \\ h^R &= \arctan\left(\frac{b^R}{a^R}\right). \end{aligned} \quad (24)$$

Among them, the index σ is the surrounding environment parameter, which is 1/2.3 in the average environment; 1/2.9 in the dark environment; 1/3.5 in the dark environment.

The LLAB model adopts the Bradford color adaptation conversion method, and the specific implementation steps are as follows.

- (1) The algorithm calculates the RGB cone response value

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_B \begin{bmatrix} \frac{X}{Y} \\ \frac{Y}{Y} \\ \frac{Z}{Y} \end{bmatrix}, \quad (25)$$

$$M_B = \begin{pmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{pmatrix}.$$

Similarly, the RGB pheasant response values of the source adaptive field white point and the reference standard white point can be calculated, which are denoted as $R_w G_w B_w$ and $R_{wr} G_{wr} B_{wr}$, respectively. Among them, the white point under the reference standard condition is CIED65, and its tristimulus value $X_{wr} Y_{wr} Z_{wr}$ is a constant ($X_{wr} = 95.05$, $Y_{wr} = 100.0$, $Z_{wr} = 108.88$).

- (2) The algorithm performs color adaptation conversion and calculates the corresponding color $X_{\text{ref}} Y_{\text{ref}} Z_{\text{ref}}$ under the reference standard conditions. The algorithm first calculates the pheasant response value $R_{\text{ref}} G_{\text{ref}} B_{\text{ref}}$ under the reference standard condition, which is

$$\begin{aligned} R_{\text{ref}} &= \left[D \left(\frac{R_{wr}}{R_w} \right) + 1 - D \right] R, \\ G_{\text{ref}} &= \left[D \left(\frac{G_{wr}}{G_w} \right) + 1 - D \right] G. \end{aligned} \quad (26)$$

If $B \geq 0$, then

$$B_{\text{ref}} = \left[D \left(\frac{B_{wr}}{B_w^p} \right) + 1 - D \right] B^p. \quad (27)$$

Otherwise,

$$B_{\text{ref}} = - \left[D \left(\frac{B_{wr}}{B_w^p} \right) + 1 - D \right] |B|^p. \quad (28)$$

Among them,

$$p = \left(\frac{B_w}{B_{wr}} \right)^{0.0834}. \quad (29)$$

Then, the algorithm calculates the corresponding color $X_{\text{ref}} Y_{\text{ref}} Z_{\text{ref}}$ under the reference standard condition, which is

$$\begin{bmatrix} X_{\text{ref}} \\ Y_{\text{ref}} \\ Z_{\text{ref}} \end{bmatrix} = M_B^{-1} \begin{bmatrix} R_{\text{ref}}Y \\ G_{\text{ref}}Y \\ B_{\text{ref}}Y \end{bmatrix}. \quad (30)$$

(3) The algorithm calculates the color appearance attributes.

The related color appearance attribute parameters and calculation methods are as follows:

$$\begin{aligned} L^L &= 116f\left(\frac{Y_{\text{ref}}}{Y_{\text{wr}}}\right)^z - 16, \\ a &= 500\left[f\left(\frac{X_{\text{ref}}}{X_{\text{wr}}}\right) - f\left(\frac{Y_{\text{ref}}}{Y_{\text{wr}}}\right)\right], \\ b &= 200\left[f\left(\frac{Y_{\text{ref}}}{Y_{\text{wr}}}\right) - f\left(\frac{Z_{\text{ref}}}{Z_{\text{wr}}}\right)\right], \\ z &= 1 + F_L\left(\frac{Y_b}{100}\right)^{(1/2)}. \end{aligned} \quad (31)$$

In the formula, the function $f(x)$ is defined as follows:

$$f(x) = \begin{cases} x^{1/F_s}, & x > 0.008856, \\ \left[\frac{(0.008856)^{1/F_s} - 16/116}{0.008856}\right]x + \frac{16}{116}, & x \leq 0.008856, \\ C^L = 25 \ln(1 + 0.05C), \end{cases}$$

$$\begin{aligned} C^L &= 25 \ln(1 + 0.05C), \\ C &= \sqrt{(a)^2 + (b)^2}, \\ h^L &= \tan^{-1}\left(\frac{b}{a}\right), \\ A^L &= C^L \cos(h^L), \\ B^L &= C^L \sin(h^L). \end{aligned} \quad (32)$$

In addition to lightness L^L , chroma C^L , and hue h^L , the model can also predict visual chroma, saturation, etc. Since these color appearance attributes are beyond the scope of this study, their specific formulas are not given.

The spectral sharpening space conversion adopts the equal energy balance matrix to convert the CIEXYZ tristimulus values to the CAT02 space, so as to obtain the sharpening space response values RGB, $X_w Y_w Z_w$ are also converted into $R_w G_w B_w$, and the specific calculation formula is as follows:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M_{\text{CAT02}} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (33)$$

$$M_{\text{CAT02}} = \begin{pmatrix} 0.7328 & 0.4296 & -0.1624 \\ -0.7036 & 1.6975 & 0.0061 \\ 0.0030 & 0.00136 & 0.9834 \end{pmatrix}.$$

The adaptive sharpening signal $R_c G_c B_c$ is obtained by multiplying the RGB signal by the independent gain control coefficient ($\alpha_R, \alpha_G, \alpha_B$), namely:

$$\begin{bmatrix} R_c \\ G_c \\ B_c \end{bmatrix} = A \bullet \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (34)$$

$$A = \begin{bmatrix} \alpha_R & 0.0 & 0.0 \\ 0.0 & \alpha_G & 0.0 \\ 0.0 & 0.0 & \alpha_B \end{bmatrix}.$$

Among them,

$$\begin{aligned} \alpha_R &= \frac{Y_w D}{R_w} + (1 - D) \\ D &= F \cdot \left[1 - \left(\frac{1}{3.6}\right)e^{(-L_A - 42/92)}\right]. \end{aligned} \quad (35)$$

The D factor is used to reflect the degree of color adaptation and is a function of the environment and L_A . Theoretically, the value of D should be from 0 (can not adapt to the selected white point) to 1 (complete adaptation). In fact, the minimum value of D does not fall below 0.65 in the dark environment and grows exponentially with L_A . The relationship curves of $D - L_A$ under the three environments are shown in Figure 5.

The cone response space transformation transforms $R_c G_c B_c$ into the pheasant response space before the nonlinear response is compressed. Through experimental comparison, it is found that nonlinear response compression in a space similar to the cone response can predict better color appearance attribute values. The conversion formula is as follows:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = M_H \cdot M_{\text{CAT02}}^{-1} \cdot \begin{bmatrix} R_c \\ G_c \\ B_c \end{bmatrix}, \quad (36)$$

$$M_H = \begin{bmatrix} 0.38971 & 0.68898 & -0.07868 \\ -0.22981 & 1.18340 & 0.04641 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}.$$

The specific formula of nonlinear response compression is as follows:

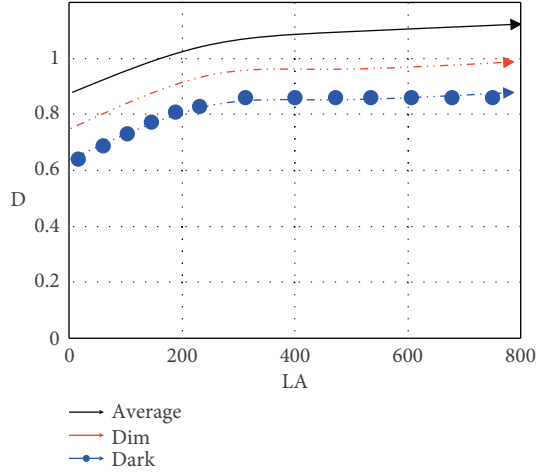


FIGURE 5: The relationship of $D - L_A$ in three environments.

$$R'_a = \frac{400(F_L R'/100)^{0.42}}{27.13 + (F_L R'/100)^{0.42}} + 0.1. \quad (37)$$

Among them,

$$F_L = 0.2k^4(5L_A) + 0.1(1 - k^4)^2(5L_A)^{1/3}, \quad (38)$$

$$k = \frac{1}{(5L_A + 1)}.$$

G'_a and B'_a are calculated similarly, and $R'_{wa} \Delta G'_{wa}$ and B'_{wa} can also be converted.

The calculation method of color appearance attribute is as follows:

$$a = \frac{R'_a - 12G'_a}{11} + \frac{B'_a}{11},$$

$$b = \left(\frac{1}{9}\right)(R'_a + G'_a - 2B'_a),$$

$$t = \frac{e^{(a^2 + b^2)^{(1/2)}}}{R'_a + G'_a + (21/20)B'_a}, \quad (39)$$

$$h = \tan^{-1}\left(\frac{b}{a}\right),$$

$$A = \left[2R'_a + G'_a + \left(\frac{1}{20}\right)B'_a - 0.305\right]N_{bb}.$$

Among them,

$$N_{bb} = 0.725\left(\frac{1}{n}\right)^{0.2}, \quad (40)$$

$$n = \frac{Y_b}{Y_w}.$$

A represents the achromatic response, and the constant term -0.305 in the formula determines the minimum luminance value, so that when Y is 0, A is also 0, and the

achromatic response A_w of the white point can be calculated similarly. The formula for calculating lightness J is

$$J = 100\left(\frac{A}{A_w}\right)^{cz}. \quad (41)$$

Among them,

$$z = 1.48 + \sqrt{n}. \quad (42)$$

With the lightness and temporary variables t , the chroma C can be calculated as follows:

$$C = t^{0.9} \sqrt{\frac{J}{100}} (1.64 - 0.29^n)^{0.73}. \quad (43)$$

4. Accurate Recognition of Motion Patterns Based on Artificial Visual Neural Network

Figure 6(a) is a schematic diagram of the visual movement of the RS neuron activated by the moving target. In this figure, a moving target (that is, the solid-line sphere in the figure) rotates CCW around the center point. At different moments in the rotational motion (i.e., the dotted sphere in the figure), the translational motion direction of the moving target changes regularly. That is, relative to the previous moment, the translational movement direction of the moving object at the current moment always changes incrementally in the counterclockwise direction. Figure 6(b) is a schematic diagram of the basic visual mechanism for RS neurons to recognize rotational motion patterns. Moreover, each step in the figure corresponds to a specific visual signal processing in the visual channel.

Figure 7(a) presents a schematic diagram of the basic visual mechanism of DRS neurons in perceiving deep rotational motion patterns. In order to perceive deep rotational motion patterns of moving objects, visual signals are processed in four steps in the visual pathway of DRS neurons, that is, (1) visual signal processing and perceived lumen change, (2) acquisition of changes in visual excitation and inhibition, (3) extraction of changes in translation direction of moving objects and depth of motion behavior, (4) synthetic depth of motion behavior and changes in translation direction, and DRS neuron output membrane potential. Each step corresponds to a specific visual signal processing process in the visual pathway. After layer-by-layer visual information processing, DRS neurons can perceive the spatiotemporal energy changes caused by deep rotational motion in the visual field. Figure 7(b) is a schematic diagram of the amount of visual motion behavior implied by a moving target that performs a counterclockwise deep rotation motion on the horizontal plane.

On the basis of the above research, the effect of the motion pattern recognition based on artificial visual neural network proposed in this paper is verified, and the motion pattern recognition effect of the model in this paper is counted, and Table 1 and Figure 8 are obtained.

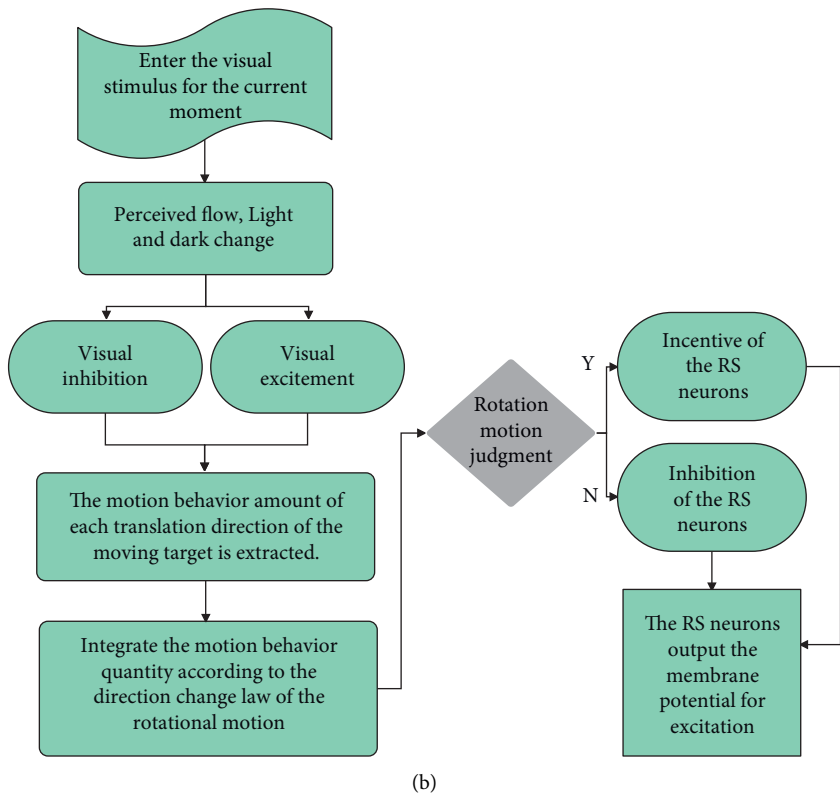
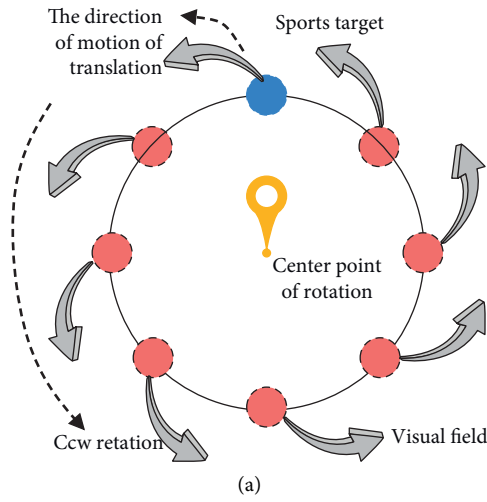


FIGURE 6: (a) Schematic diagram of the movement direction change of the moving target in the frontal-parallel plane rotating motion; (b) the basic visual mechanism of RS neuron detecting the rotational movement pattern.

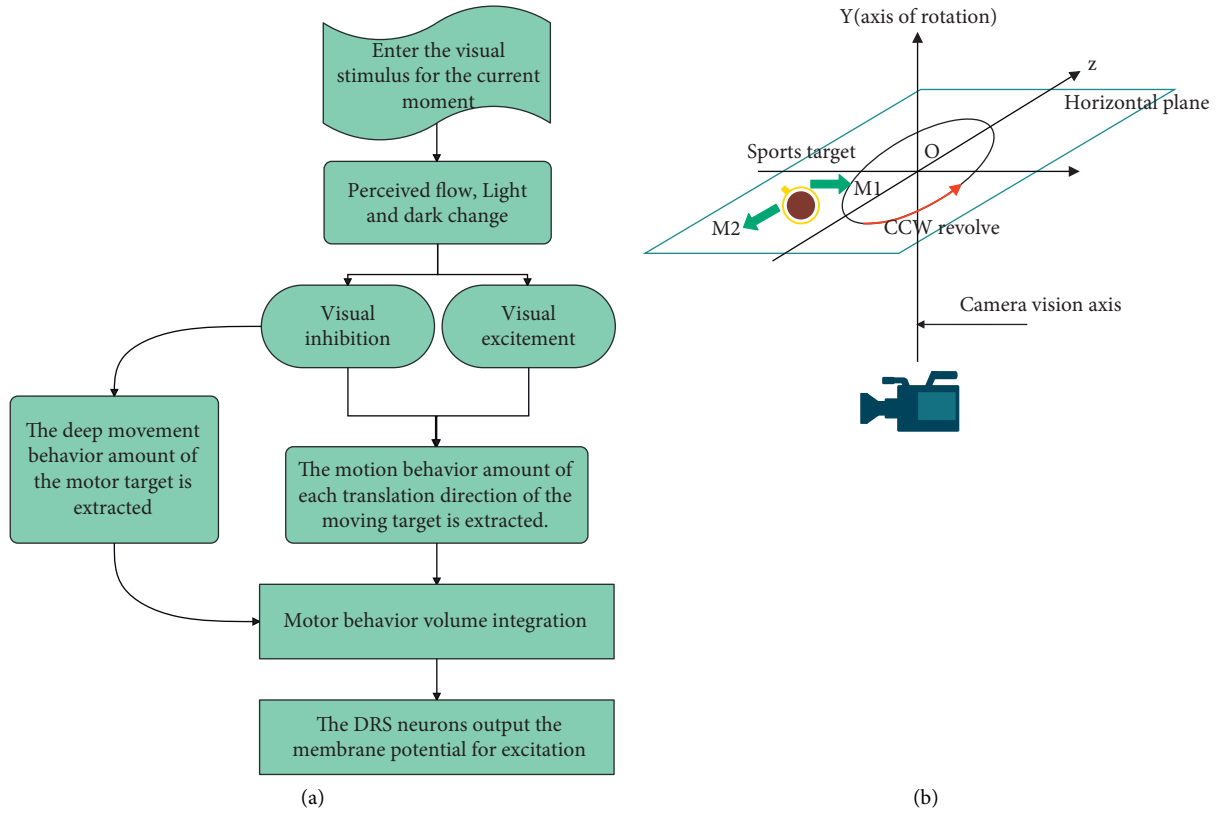


FIGURE 7: (a) Schematic diagram of the basic visual mechanism of deep rotational motion perception by DRS neurons. (b) Schematic diagram of the hidden visual motion behavior of a moving target that performs counterclockwise depth rotation motion on the horizontal plane. $M1$ in the figure represents the translation direction, and $M2$ represents the depth motion direction.

TABLE 1: Verification of recognition effect of motion pattern recognition model based on artificial visual neural network.

Num	Motion recognition	Num	Motion recognition	Num	Motion recognition
1	89.30	21	81.08	41	78.75
2	79.09	22	78.60	42	83.52
3	78.88	23	79.49	43	85.55
4	81.69	24	78.93	44	80.82
5	86.81	25	82.57	45	81.64
6	88.32	26	86.38	46	81.79
7	80.60	27	89.53	47	77.45
8	82.36	28	89.12	48	80.28
9	80.61	29	89.57	49	82.47
10	86.12	30	82.78	50	79.96
11	82.61	31	86.05	51	82.02
12	83.85	32	78.75	52	79.81
13	87.45	33	86.51	53	84.38
14	79.51	34	90.02	54	81.15
15	87.51	35	86.79	55	85.63
16	83.71	36	81.96	56	83.33
17	87.69	37	79.60	57	83.18
18	79.74	38	87.92	58	82.52
19	82.89	39	81.10	59	89.61
20	80.33	40	78.92	60	86.18

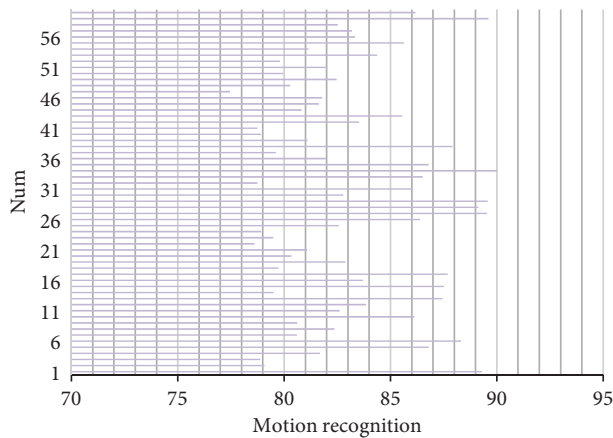


FIGURE 8: Statistical diagram of the recognition effect of the motion pattern recognition model based on artificial visual neural network.

Combined with the simulation recognition experiment, it can be seen that the motion pattern recognition model based on artificial visual neural network can accurately identify the motion pattern category.

5. Conclusion

The motion pattern recognition of moving objects is generally divided into training process and testing process. In the training process, all training trajectories (trajectories of known motion patterns) are preprocessed first, then features are extracted for each trajectory, and finally a classifier is constructed. In the test phase, all test trajectories (trajectories with unknown motion patterns) are preprocessed first, then features are extracted for each test trajectory, and finally the motion pattern recognition is performed on the test trajectories using the classifier obtained in the training phase. This paper combines the artificial visual neural network to construct a motion pattern recognition system to improve the motion pattern recognition effect. The experimental results show that the motion pattern recognition system based on artificial visual neural network can accurately identify the motion pattern category.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This study was sponsored by Hubei University of Technology.

References

- [1] G. Szűcs and B. Tamás, "Body part extraction and pose estimation method in rowing videos," *Journal of Computing and Information Technology*, vol. 26, no. 1, pp. 29–43, 2018.
- [2] R. Gu, G. Wang, Z. Jiang, and J. N. Hwang, "Multi-person hierarchical 3d pose estimation in natural videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 11, pp. 4245–4257, 2019.
- [3] M. Nasr, H. Ayman, N. Ebrahim, R. Osama, N. Mosaad, and A. Mounir, "Realtime multi-person 2D pose estimation," *International Journal of Advanced Networking and Applications*, vol. 11, no. 6, pp. 4501–4508, 2020.
- [4] N. T. Thành and P. T. Công, "An evaluation of pose estimation in video of traditional martial arts presentation," *Journal of Research and Development on Information and Communication Technology*, vol. 2019, no. 2, p. 114, 2019.
- [5] I. Petrov, V. Shakhuro, and A. Konushin, "Deep probabilistic human pose estimation," *IET Computer Vision*, vol. 12, no. 5, pp. 578–585, 2018.
- [6] G. Hua, L. Li, and S. Liu, "Multipath affinity stacked—hourglass networks for human pose estimation," *Frontiers of Computer Science*, vol. 14, no. 4, pp. 1–12, 2020.
- [7] K. Aso, D. H. Hwang, and H. Koike, "Portable 3D human pose estimation for human-human interaction using a chest-mounted fisheye camera," in *Proceedings of the Augmented Humans Conference 2021*, pp. 116–120, Geneva, Switzerland, February 2021.
- [8] D. Mehta, S. Sridhar, O. Sotnychenko et al., "VNect," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–14, 2017.
- [9] S. Liu, Y. Li, and G. Hua, "Human pose estimation in video via structured space learning and halfway temporal evaluation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 7, pp. 2029–2038, 2018.
- [10] S. Ershadi-Nasab, E. Noury, S. Kasaei, and E. Sanaei, "Multiple human 3d pose estimation from multiview images," *Multimedia Tools and Applications*, vol. 77, no. 12, Article ID 15601, 2018.
- [11] X. Nie, J. Feng, J. Xing, S. Xiao, and S. Yan, "Hierarchical contextual refinement networks for human pose estimation," *IEEE Transactions on Image Processing*, vol. 28, no. 2, pp. 924–936, 2018.
- [12] Y. Nie, J. Lee, S. Yoon, and D. S. Park, "A multi-stage convolution machine with scaling and dilation for human pose estimation," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 13, no. 6, pp. 3182–3198, 2019.
- [13] A. Zarkeshev and C. Csiszár, "Rescue method based on V2X communication and human pose estimation," *Periodica Polytechnica: Civil Engineering*, vol. 63, no. 4, pp. 1139–1146, 2019.
- [14] W. McNally, A. Wong, and J. McPhee, "Action recognition using deep convolutional neural networks and compressed spatio-temporal pose encodings," *Journal of Computational Vision and Imaging Systems*, vol. 4, no. 1, p. 3, 2018.
- [15] R. G. Díaz, F. Laamarti, and A. El Saddik, "DTCoach: your digital twin coach on the edge during COVID-19 and beyond," *IEEE Instrumentation and Measurement Magazine*, vol. 24, no. 6, pp. 22–28, 2021.
- [16] A. Bakshi, D. Sheikh, Y. Ansari, C. Sharma, and H. Naik, "Pose estimate based yoga instructor," *International Journal of*

- Recent Advances in Multidisciplinary Topics*, vol. 2, no. 2, pp. 70–73, 2021.
- [17] S. L. Colyer, M. Evans, D. P. Cosker, and A. I. T. Salo, “A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system,” *Sports medicine - open*, vol. 4, no. 1, pp. 24–15, 2018.
- [18] I. Sáráandi, T. Linder, K. O. Arras, and B. Leibe, “Metrabs: metric-scale truncation-robust heatmaps for absolute 3d human pose estimation,” *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 3, no. 1, pp. 16–30, 2020.
- [19] A. Azhand, S. Rabe, S. Müller, I. Sattler, and A. Heimann-Steinert, “Algorithm based on one monocular video delivers highly valid and reliable gait parameters,” *Scientific Reports*, vol. 11, no. 1, Article ID 14110, 2021.
- [20] J. Bhombe, A. Jethwa, A. Singh, and T. Nagarhalli, “Review of pose recognition systems,” *VIVA-Tech International Journal for Research and Innovation*, vol. 1, no. 4, pp. 1–8, 2021.
- [21] D. P. P. Nagalakshmi Vallabhaneni, “The analysis of the impact of yoga on healthcare and conventional strategies for human pose recognition,” *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 6, pp. 1772–1783, 2021.
- [22] J.-J. Liu, J. Newman, and D.-J. Lee, “Using artificial intelligence to provide visual feedback for golf swing training,” *Electronic Imaging*, vol. 33, no. 6, pp. 321–1, 2021.
- [23] D. Yang, “Research on multi-target tracking technology based on machine vision,” *Applied Nanoscience*, vol. 2, pp. 1–11, 2022.