www.cad-journal.net
Computer-Aided Design AND APPLICATIONS

Music Rhythm Recognition and Dynamic Design Based on Computer Vision

Fan Yang

Huanghe Jiaotong University, Wuzhi, Henan 454950, China,
2020021618@zjtu.edu.cn

Corresponding author: Fan Yang, 2020021618@zjtu.edu.cn

Abstract. As a powerful image processing and understanding technology, computer vision provides a new perspective and method for research and application in the fields of music and Computer-aided design (CAD). This paper mainly discusses the application of computer vision in music rhythm recognition and CAD dynamic design. By combining computer vision with Deep learning (DL) technology, this paper proposes a music rhythm recognition method based on a Convolutional neural network (CNN) and verifies its effectiveness in different types of music videos. The results show that the music rhythm recognition method based on CNN proposed in this paper has certain effects on different music types. It performs best, especially in pop music, but it also maintains a relatively high recognition rate in classical music and folk music, which can basically reach more than 90% accuracy. This verifies the generalization ability and effectiveness of this method. At the same time, the method in this paper can reflect the change of music rhythm to CAD design in real-time and realize the dynamic and interactive design. The research results show the unique application and value of computer vision technology in the fields of music and CAD, which is expected to inspire and promote research and practice in related fields.

Keywords: Computer Vision; Music Rhythm Recognition; Computer-Aided Design; Deep Learning; Convolutional Neural Network
DOI: https://doi.org/10.14733/cadaps.2024.S15.227-243

1 INTRODUCTION

Music recognition is an important research direction in the field of artificial intelligence, which involves the processing and understanding of audio signals. Although many methods have been used for music recognition, optical music recognition methods have their unique advantages. This method can directly convert music score images into readable music information without the need for audio signal conversion. Alfaro et al. [1] explored how to utilize the two-dimensional properties of unknown music symbols for neurooptical music recognition. Unknown musical symbols are a special type of symbol with two-dimensional properties. This symbol is composed of a series of points and lines, which can represent various elements in music, such as notes, rhythm lines, and chords. Due to the two-dimensional nature of this symbol, we can use image-processing techniques to analyze and recognize it. In neurological music recognition, we first need to preprocess the score image to extract unknown music symbols from it. Then, we can use neural network technology to classify and recognize these symbols. Computer vision involves many disciplines, such as image processing, machine learning, pattern recognition, and so on, and its application scenarios are wide, including but not limited to robot navigation, autonomous driving, medical image analysis, security monitoring, and so on. With the progress and innovation of education, it is particularly important for teachers to prepare for the initial curriculum. This is especially true for new creative arts and design courses. In these courses, visual art options are an important component, which not only requires students to master basic artistic skills but also requires them to be able to understand and apply the basic principles of visual art. Alhassan et al. [2] explored its performance in visual arts options. Through comparative research, we can conclude that JHS schools perform well, especially in terms of visual arts options. However, we can also see that other schools perform better in certain aspects. For example, some schools place more emphasis on incorporating more technical elements into their curriculum, while others place more emphasis on cultivating students' creativity and critical thinking. Therefore, we suggest that teachers at JHS schools should consider borrowing from the practices of other schools to improve their curriculum further. Teachers can consider adding some technical elements to the course, such as digital painting, 3D modeling, etc., to help students master more artistic skills. Teachers can design activities that require students to unleash their creativity, such as creating their own artwork and designing their own clothing. With the rapid development of science and technology, computer vision technology has gradually penetrated into all aspects of people's lives. Computer vision technology can realize the automatic processing and understanding of visual information such as images and videos, and provide new solutions and methods for many fields. Optical music recognition is a technology that converts music score images into readable music information. This technology involves identifying and understanding various elements in music score images, such as notes, rhythm lines, and chords. Among them, the recognition of note position is an important link in optical music recognition. Accurately identifying the position of notes can help to interpret music scores more accurately, thereby helping musicologists, performers, and machine learning researchers improve and optimize related technologies for music learning and performance. The application of convolutional neural networks (CNN) in image processing and computer vision has made significant progress. Especially in optical music recognition, CNN has been proven to be an effective technique for processing and analyzing various elements in music score images. Andrea and Zahra [3] explore how to use CNN to identify the position of notes in music score images. A database containing multiple scores was used for the experiment, and the results showed that using CNN to identify the position of notes in the score image is feasible. By adding an additional layer to the CNN model to process note position information. We also found that attention mechanisms and techniques such as CRF can further improve the accuracy of note position recognition.

In the era of multimedia information, interaction design plays an important role in various fields, especially in the fields of music and design. Music is one of the important ways for humans to express emotions and experience beauty, while design is an important means for humans to create and change the world. Ba et al. [4] explored how to use interaction methods based on speech perception and visual images to achieve the combination of music rhythm recognition and CAD dynamic design, providing users with a more intuitive, flexible, and efficient interaction experience. Speech perception technology is a technology that utilizes computers to recognize and process human speech signals. In the field of music, speech perception technology can be used to recognize the rhythm of music. Specifically, by extracting features and pattern recognition from speech signals, human speech signals can be transformed into computer-recognized music rhythm signals. Music recognition is an important application in the field of computer vision, which can be used for various tasks, such as automatic music arrangement, music classification, music search, etc. Among various music recognition tasks, handwritten music recognition and optical music recognition are the two main forms. Handwritten music recognition mainly focuses on identifying notes and their connections from handwritten music scores on paper or tablets, while optical music recognition mainly focuses on identifying notes and their connections from music scores printed in books, magazines, or newspapers. Baro et al. [5] explored the foundations and challenges of these two music recognition tasks and described a baseline model that can be applied to these two tasks. Handwritten music recognition is an important application in the field of computer vision, with the goal of identifying notes and their connections from handwritten music scores. The challenge of handwritten music recognition lies in the diversity and uncertainty of handwritten music scores. People's handwriting styles often vary greatly, and the relationships between notes in handwritten music scores are often difficult to recognize accurately. In addition, there may be various errors and blurriness in handwritten music scores, which can also increase the difficulty of recognition. Music, as a universal art form, its rhythm is one of the core elements of music. Traditional music rhythm recognition methods are often based on audio signals, but this method may not be able to fully extract rhythm information for some scenes, such as music videos and music-related visual arts. Therefore, identifying music rhythm from the perspective of computer vision has become a problem worthy of study.

In our daily lives, music is everywhere, and its unique charm influences our emotions, thinking, and behavior. However, for those with impaired vision or inability to read music directly, understanding and appreciating music can be a challenge. This is the application scenario of optical music recognition, which transforms the visual information of music into perceivable sound through technological means, providing a way for visually impaired individuals to perceive and appreciate music. Calvo et al. [6] explore the technical principles, application fields, and development prospects of optical music recognition. For people with impaired vision or inability to directly read music scores, optical music recognition technology can convert music scores into sound, helping them better understand and appreciate music. Optical music recognition technology can help students improve their solfeggio and ear training skills. In addition, it can also help teachers evaluate students' music perception ability. Optical music recognition technology can help researchers analyze the structure and features of music scores. In addition, it can also be used to develop new forms of musical expression and instruments. In the ocean of musical expression, body language and gestures are often used as expressive expressions. From ancient dances to modern stage plays, gestures play an important role in performing arts. In the field of computer music, gestures are also widely used, not only as a part of performance but also as an important tool for creating and interpreting music. Gesture language and music have a deep connection. Since ancient times, humans have used gestures to convey information and emotions, especially in the field of music. For example, in classical music, the conductor uses complex gestures to guide the orchestra's performance, conveying the emotions and structure of the music to every performer. In folk music, dancers create infectious performances by combining body movements and gestures with melody and rhythm. With the development of computer music, gestures, as an input device, are increasingly widely used in music creation and performance. For example, in electronic music, performers often use touch screens, controllers, or specialized sensors to generate sound. These devices can capture the gestures and movements of performers and convert them into music signals. Cavdir et al. [7] further expanded the expression possibilities of music through this approach. Although the current CAD design has made some automation achievements, the traditional CAD design mainly depends on the designer's experience and manual operation. For complex design tasks that need dynamic adjustment, the traditional methods are inadequate. Therefore, how to use computer vision technology to realize the dynamic design of CAD and improve design efficiency and quality is also a subject with research value. Against this background, this paper introduces computer vision technology into music rhythm recognition and CAD dynamic design, trying to explore new processing methods and application possibilities for these two fields. The innovations of this paper mainly include:

㊀ Traditional music rhythm recognition is mainly based on audio signals, but this paper innovatively proposes to use computer vision technology to extract visual features from music videos for rhythm recognition. Analyzing the visual elements in music videos allows us to realize more comprehensive and accurate rhythm recognition.

㊁ Traditional CAD design mainly depends on the designer's experience and manual operation, but this paper explores the possibility of using computer vision technology for dynamic CAD design. The information from design sketches can be automatically extracted and transformed into CAD models through image processing and recognition. This is an innovative application in the CAD field.

⊛ This paper not only independently studies music rhythm recognition and CAD dynamic design, but also integrates them. By applying the results of music rhythm recognition to CAD dynamic design, the cross-modal interaction from music to design can be realized, which provides more possibilities for creative design and artistic expression. This integration method is seldom explored in previous studies, so it constitutes an innovation of this paper.

Starting from the introduction, this paper gradually goes deep into the computer vision foundation, music rhythm recognition principle, CAD dynamic design principle, and other aspects. Finally, the system integration and simulation research are carried out to verify the effectiveness of the proposed method. Each section will be based on the research in the previous section to ensure the logic and coherence of the article.

2 RELATED WORK

In the field of music, optical music recognition is a technology that converts music score images into readable music information. Although significant progress has been made in this technology, there are still some challenges in processing complex and ever-changing score images. Garrido et al. [8] proposed a holistic approach from image to graphics to address this issue. This method has significant effects in analyzing international literature and identifying musical features. In optical music recognition, the overall method of image-to-graphics is an effective technique. By combining image processing with graphics technology, we can better understand and recognize various elements and features in music scores. In addition, this method can also help us analyze and identify music scores from different languages and cultures, thereby promoting the exchange and inheritance of music culture. In the future, further research and improvement will be conducted on the overall method of image-to-graphics to improve its application in optical music recognition. At the same time, we also hope to see more research and practical exploration on how to apply this technology to fields such as education and cultural inheritance. Music feature analysis is the key to automatic music classification. By extracting and analyzing the basic features of music, such as pitch, rhythm, harmony, etc., Ge et al. [9] effectively classified and recognized music. In addition, feature analysis can also help us understand and compare the similarities and differences between different music, thereby providing support for music recommendation and search applications. In order to extract music features more effectively, we need to improve existing feature extraction algorithms. For example, the accuracy of feature extraction can be improved by increasing sensitivity to features such as pitch, rhythm, and harmony or by introducing more complex feature analysis algorithms such as deep learning algorithms. In the music classification stage, we need to classify based on the characteristics of the music automatically. In terms of system architecture, parallel computing, distributed storage, and other technologies can be used to improve the processing speed and storage capacity of the system. In addition, technologies such as cloud computing can also be introduced to achieve automatic classification of large-scale music. We can improve the classification accuracy and processing speed of music, providing users with a better music browsing and search experience. This can also provide strong support for applications such as music recommendation and marketing, promoting the development of the music industry. Computer-assisted music education and music creation have become an emerging trend in the field of music education. Maba [10] explores the advantages of computer-assisted music education and how to utilize computer-assisted music creation. Computers can use various intelligent teaching tools, such as music production software, digital audio workstations, etc. These tools can help students better master music knowledge and enhance their interest in learning. Music production software can help students better achieve music creation. They can create music works through the software, adjust parameters such as sound effects and timbre, and achieve the best music effects. Synthesizers can generate various unique timbres, helping students create unique musical works. Samplers can collect various real-life sound materials, such as natural sounds and human language, to help students create richer music works. Computer-assisted music education and music creation are the future development trends in the field of music education.

Handwritten music symbol recognition is an important research direction in the field of optical music recognition, which involves the recognition and understanding of handwritten notes, rhythm lines, chords, and other elements in music score images. Due to the diversity and complexity of handwritten music symbols, accurate recognition of these symbols requires an efficient and reliable method. Paul et al. [11] introduced a model set based on deep transfer learning for handwritten music symbol recognition. Deep transfer learning is an effective machine learning method that solves the problems of data sparsity and feature extraction in deep learning by transferring parameters from pre-trained models to new tasks. In handwritten music symbol recognition, deep transfer learning model sets are used to improve recognition accuracy and robustness. This model utilizes a pre-trained recurrent neural network (RNN) model to model and recognize handwritten music symbol sequences. By combining pre-trained RNN models with new datasets, accurate recognition of handwritten music symbols can be achieved. It fully utilizes the interactivity of computers, information storage and retrieval capabilities, and the richness of network resources, improving teaching quality and efficiency. Especially in music appreciation courses, the application of CAI has changed the traditional teaching mode, providing learners with richer learning resources and a more intuitive learning experience. The music resources on the internet are extremely rich, including various types of music, such as classical music, pop music, and ethnic music, as well as various styles of music, such as jazz and electronic music. These resources provide more options for music appreciation courses, meeting the interests and needs of different students. The music resources on the internet have high interactivity. Students can listen, download, and share music online, as well as participate in music creation and adaptation. This interactivity not only enhances students' learning enthusiasm but also improves their learning effectiveness. Network resources make remote access possible. No matter where they are, as long as there is an internet connection, students can access music resources anytime and anywhere for autonomous learning. Pei et al. [12] utilized the diversity and interactivity of network resources, as well as the standardization and efficiency of computer-aided teaching management systems.

Optical music recognition is a technology that converts music score images into readable music information. However, due to differences in different languages and cultures, there are still some challenges in the encoding steps of optical music recognition. To solve this problem, we can apply automatic translation technology to the encoding steps of optical music recognition. Ríos et al. [13] explore how to apply automatic translation technology to the encoding steps of optical music recognition. Before inputting music score images into an optical music recognition system, we can use automatic translation technology to preprocess the images. For example, we can use techniques such as image segmentation and character recognition to separate various elements in the image, such as notes, rhythm lines, and chords, and identify their types and positions. In the encoding step of optical music recognition, we need to extract various features from the music score image, such as the shape, size, position, etc., of notes. We can use automatic translation technology to classify and recognize these features, thereby obtaining more accurate results. In the encoding steps of optical music recognition, we can use automatic translation technology to optimize the encoding algorithm. Thereby improving the efficiency and accuracy of coding. In music theory, rhythm is one of the core components of music, shaping the structure and emotions of music through the organization and allocation of time. In cyclic rhythm mode recording, the repetition and variation patterns of music rhythm are more prominent. Rocamora et al. [14] explored how to use information theory concepts to analyze the musical rhythm of cyclic rhythm mode recordings in order to reveal its internal structure and characteristics. Information theory is a discipline that studies the measurement, transmission, and transformation laws of information, mainly involving issues such as encoding, storage, transmission, and extraction of information. In music analysis, information theory provides an effective tool for measuring and analyzing the amount of information in music works, revealing their structure and characteristics. In cyclic rhythm mode recording, the repetition and regularity of music information are strong. We can measure the complexity and degree of variation of music rhythm by calculating the amount of information in repetitive patterns. In music rhythm analysis, we can apply the concepts of information encoding and decoding to transform music rhythm into a form that is easy to analyze and understand. For example, music rhythms can be converted into binary code or note sequences for subsequent processing and analysis. In cyclic rhythm mode recording, channel capacity can be used to measure the amount of information transmitted by music rhythm. By analyzing channel capacity, we can understand the limitations and characteristics of music rhythm in the transmission process.

Computer-aided design has also demonstrated its enormous potential. By using a 3D-aided design system, dance designers can design dance movements more intuitively and accurately, thereby improving the artistic expression and appreciation of dance. Tan and Yang [15] explore how to design dance movements based on a computer-assisted three-dimensional system. Through motion capture technology, the movements of dancers can be converted into digital information and input into a computer three-dimensional assistance system. This allows designers to accurately obtain data on dance movements, including amplitude, speed, angle, etc. After obtaining the data on dance movements, the computer three-dimensional assistance system can perform action editing. Designers can modify, optimize, and even innovate dance movements through the system. This greatly improves the efficiency and flexibility of dance action design. Computer three-dimensional assistance systems can generate. In addition, the system can also perform simulation demonstrations, allowing designers to preview and adjust dance movements before actual performances. Multi-font digital score recognition is an important research direction in the field of optical music recognition, which involves the recognition and understanding of different fonts and styles of numbers in score images. Due to significant differences in shape, size, and arrangement among numbers of different fonts and styles, accurately identifying these numbers is a challenging task. Wang et al. [16] introduced a model based on kernel density estimation and convolutional neural networks for multi-font digital score recognition. Kernel density estimation is a nonparametric statistical method that can be used to estimate numerical solutions of probability density functions. In multi-font digital score recognition, we can use kernel density estimation to extract the shape features of numbers. By applying convolutional neural networks to multi-font digital score recognition, we can achieve accurate recognition and understanding of handwritten and printed numbers with multiple fonts and styles. To verify the performance of the proposed model based on kernel density estimation and convolutional neural network in multi-font digital score recognition, we conducted experiments using a digital score dataset containing multiple fonts and styles. The experimental results show that the proposed model can achieve accurate and efficient number recognition and has good robustness for handwritten and printed numbers with multiple fonts and styles. Meanwhile, kernel density estimation has also achieved good results in extracting digital shape features.

With the deepening development of globalization, the position of English translation teaching in higher education is becoming increasingly important. Traditional translation teaching methods often focus on theoretical teaching while neglecting practical aspects, which makes it difficult for students to master translation skills truly. In recent years, the rapid development of computer-aided translation software (CAT) has provided new ideas and methods for English translation teaching. By constructing a parallel corpus, CAT software can be effectively introduced into translation teaching, improving teaching quality and efficiency. Wang et al. [17] explored how to construct. In English translation teaching, parallel sentence pairs can provide students with references during the translation process, helping them better understand the conversion rules and techniques between the source language and the target language. In addition, by analyzing parallel sentence pairs, students can also learn cultural differences and expressions between different languages and improve cross-cultural communication skills. By adjusting the parameters and algorithms of CAT software, the performance of the corpus is optimized to improve translation accuracy and fluency. Apply the constructed parallel corpus to English translation teaching to assist students in translation practice and autonomous learning. With the rapid development of the music industry, the demand for music recognition technology is growing day by day. Especially in the field of piano music, the recognition of piano music is more challenging due to its complex performance techniques and rich musical expression. To address this issue, we can use market fingerprint technology for large-scale multimodal piano music recognition. Yang et al. [18] explore how to apply market fingerprint technology to large-scale multimodal piano music recognition. By extracting the characteristics of piano timbre and utilizing market fingerprint technology, different timbres are classified and recognized. Different performance styles are classified and recognized by extracting features of piano performance styles and utilizing market fingerprint technology. By extracting emotional features from piano music, market fingerprint technology is used to classify and recognize different emotions. To verify the performance of the proposed market fingerprint technology in large-scale multimodal piano music recognition, we conducted experiments using a piano music dataset containing multiple timbres, performance styles, and emotional labels. The experimental results show that the proposed market fingerprint technology can achieve accurate and efficient piano music recognition, and has good processing performance for large-scale multimodal piano music datasets. It can simulate real instrument sounds, provide rich music resources, and intuitively display complex music theories to students through forms such as graphics and animations. Yuan et al. [19] explored the requirements of system design, we need to choose appropriate hardware and software to implement the system. The computer-assisted music teaching system based on interactive mode is an advanced music teaching method. It provides rich music resources, as well as intuitive graphical and animated displays, by simulating real instrument sounds and music theory, enabling students to have a deeper understanding and mastery of music knowledge. Meanwhile, through interactive functions, teachers can more effectively manage and monitor students' learning progress and effectiveness. However, the design and implementation of the system require detailed requirements analysis, architecture design, module design, database design, and implementation processes. After completing the design and implementation of the system, strict testing and optimization are also required to ensure the stability and functionality of the system.

3 PRINCIPLE AND METHOD OF MUSIC RHYTHM RECOGNITION

3.1 Basic Concepts of Music Rhythm

In the field of music, computer vision can be used for music video analysis, music performance recognition, and other tasks. For example, by analyzing the visual elements in music videos, we can recognize and understand the rhythm and emotion of music. In addition, computer vision can also be used for automatic reading and analysis of music scores and automatic evaluation of music performances. In the field of CAD, computer vision is mainly used in aided design and automatic modeling. Through the technology of image processing and recognition, information such as lines and shapes in design sketches can be automatically extracted and transformed into CAD models. In addition, computer vision can also be used for tasks such as automatic repair and optimization of CAD models to improve design efficiency and quality. Music rhythm is one of the core elements of music, which defines the strength, length, and relationship between notes in music. Rhythm is the skeleton of music, which gives music power and rhythm. In music, rhythm is usually composed of beat, rhythm pattern, and sense of rhythm. The beat of music refers to the division of the time unit of music according to the specified time interval, which is usually expressed by the time sign, such as 4/4 beat, 3/4 beat, and so on. Rhythm is the combination of repeated notes and rests in a beat period, which determines the style and characteristics of music. The sense of rhythm refers to people's perception and cognition of the rhythm in music, which is one of the important factors for people to feel and express music.

3.2 Music Rhythm Recognition Method Based on Computer Vision

Traditional music rhythm recognition methods are mainly based on audio signals. These methods usually include three main steps: audio signal processing, feature extraction, and rhythm pattern recognition. Based on these features, machine learning or signal processing algorithm is used to identify and analyze rhythm patterns. These methods are widely used in music information retrieval, music creation, and other fields. However, the traditional audio-based music rhythm recognition method can not make full use of the visual information in music videos, and the recognition results of music performances and music videos containing rich visual elements will be limited to some extent.

The music rhythm recognition method based on computer vision recognizes music rhythm by extracting visual features from music videos. This paper adopts the CNN algorithm. First, the music data is preprocessed for input into CNN. The preprocessing step includes converting the audio signal into a spectrogram or Mel spectrogram, and these images can capture the frequency and time information of audio. Then, the music rhythm label is converted into a format that can be understood by the neural network. In this paper, the rhythm type is converted into single heat coding. Then a CNN model is constructed, which can receive preprocessed images as input. CNN usually includes multiple convolution layers, pooling layers, and fully connected layers. Assuming that the number of notes is N, a maximum likelihood model can be established:

$$\left(\hat{f}_0^1,\ldots,\hat{f}_0^N\right) = \arg\max_{f_0^1,\ldots,f_0^N \in F} p\left(O \middle| f_0^1,\ldots,f_0^N\right) \tag{1}$$

$$X_{STFT}(k,n) = \sum_{m=0}^{N-1} x(n-m)\, w(m)\, e^{-j2k\pi m/N} \tag{2}$$

$$p(k) = \left[12\log_2\left(\frac{k}{N\cdot f_{sr}}\cdot f_{ref}\right)\right] \bmod 12 \tag{3}$$

$$PCP(p) = \sum_{k=p(k)=p} \left|X(K)\right|^2 \qquad p = 1,2,3,\ldots,11 \tag{4}$$

The outline feature of the tone level represents a tone level by a vector of twelve dimensions, which reflects the relative intensity of notes under the chromatic scale in each twelve-tone interval. In this paper, at the end of the model, a softmax layer is set to output the probability of each rhythm type. If linear regression is used for the sample (x_i, y_i) $(i = 1,2,3,\ldots,m)$, its loss function is:

$$J(\theta_0,\theta_1) = \sum_{i=1}^{m} \left(h\theta(x_i) - y_i\right)^2 \tag{5}$$

Among them, x_i represents the i th sample feature; y_i represents the output corresponding to the i th sample; $h\theta(x_i)$ is the hypothesis function. The network contains m neurons corresponding to m types of music styles, and the output probability is:

$$p = \left[P_1, P_2, P_3, \ldots, P_m\right]^T \tag{6}$$

Use the Softmax regression formula as follows:

$$p_j = \frac{\exp\left(X_8^j\right)}{\sum_{i=1}^{m} \exp\left(X_8^i\right)} \qquad j = 1,2,3,\ldots,m \tag{7}$$

$$H(z) = 1 - \mu z^{-1} \tag{8}$$

The signal equation is:

$$y(n) = x(n) - \mu x(n-1) \tag{9}$$

Where μ is the pre-emphasis coefficient, which is a parameter close to 1. The μ of this system is 0.98.

Finally, based on the extracted visual features, the CNN algorithm is used to identify the music rhythm. Compared with the traditional methods, the music rhythm recognition method based on computer vision can make full use of the rich visual information in music videos and provide more comprehensive and accurate rhythm recognition results. At the same time, this method can also be applied to the automatic annotation of music videos, the evaluation of music performances, and other fields to promote the interdisciplinary research and application of music and computer vision.

4 PRINCIPLE AND METHOD OF CAD DYNAMIC DESIGN

4.1 Dynamic CAD Design Concept

Dynamic CAD design is a design method that introduces dynamics and adaptability into the CAD process. It allows designers to make real-time adjustments and modifications according to the actual situation and needs during the design process. Compared with traditional static CAD design, dynamic CAD design is more flexible and can respond in real-time according to user's needs, design constraints, and external changes. In dynamic CAD design, the key concepts include design parameterization, design constraints, and dynamic simulation. Parameterization of design refers to associating design elements with parameters through the parametric modeling method, so as to change the design by adjusting parameters. Design constraints refer to the restrictions and rules set in the design to ensure its feasibility and meet the specifications' requirements. Dynamic simulation previews the design results in real-time through computer simulation technology so that designers can make timely adjustments.

4.2 5Dynamic Design Method of CAD Based on Computer Vision

The dynamic CAD design method based on computer vision is an innovative method that applies computer vision technology to the dynamic CAD design process. It provides dynamic input and guidance for CAD design by extracting and analyzing visual features in design images. The specific method steps are shown in Table 1.

Step	*Operate*	*Describe*
Image acquisition and processing	Obtaining design images through computer vision technology	Using cameras, scanners, and other equipment to obtain design images, and pre-processing operations, such as denoising, image enhancement, etc., to extract clear image features.
	Pre-processing operations (denoising, image enhancement, etc.)	The purpose of preprocessing is to improve the image quality, highlight the features, and provide accurate input for the following steps.
Feature extraction and recognition	Using computer vision algorithm to extract design-related feature information	The edge detection algorithm is used to extract the line features in the image, and the shape recognition algorithm is used to identify the shape features, such as circles and squares, as well as to analyze the color distribution.
	Extracted features such as lines, shapes, colors	These features can describe the geometric shape, outline, and color distribution of the design and provide dynamic input and guidance for CAD models.
Parametric modeling	Associating the extracted features with the parameters of the CAD model.	By establishing the corresponding relationship between features and design parameters, the parametric representation of design elements is realized. When the image features change, the corresponding design parameters will be automatically updated.
	Establish parametric model	The parametric model allows designers to change the shape, size, and other attributes of the CAD model by adjusting parameter values, thus realizing dynamic design.
Dynamic design and simulation	Dynamic CAD design based on a parametric model	Designers can drive the dynamic updating of CAD models by adjusting image features or directly modifying parameters. This method can respond to the change in design requirements in real time and accelerate the design iteration process.

	Preview design results in real-time	Using computer simulation technology, the change results of the CAD model can be viewed in real-time during the design process. This helps designers to adjust design parameters in time, verify the feasibility of design ideas, and achieve the final design goal.

Table 1: CAD dynamic design based on computer vision.

Generally speaking, the CAD dynamic design method based on computer vision has the following advantages:

Intuitive: Designers can intuitively understand the design intent and design results by observing the image features.

Flexibility: The extraction and recognition of image features can flexibly adapt to various design requirements and changes and realize the dynamic design process.

Improvement of automation: Automatic feature extraction and recognition by computer vision technology can reduce designers' manual labor and improve design efficiency.

5 INTEGRATION OF MUSIC RHYTHM RECOGNITION AND CAD DYNAMIC DESIGN BASED ON COMPUTER VISION

The integration principle of music rhythm recognition and CAD dynamic design based on computer vision is mainly based on the idea of cross-modal interaction and multidisciplinary integration. Music rhythm recognition and CAD dynamic design are essentially processes of feature extraction and pattern recognition for specific inputs (audio and visual). Therefore, there is an inherent relationship and correspondence between them, which provides a basis for integration. The key to the integration of the system is to establish the mapping relationship between music rhythm and CAD design elements and realize the automatic establishment and dynamic adjustment of this mapping relationship through computer vision technology. In this way, the change of music rhythm can drive the dynamic adjustment of CAD design in real time, so that the design process is closely related to music rhythm. The linear MFCC coefficient extraction of audio short-time frames is shown in Figure 1.

In order to establish the corresponding relationship between music rhythm and CAD dynamic design, this paper first defines a set of mapping rules. This set of rules is based on the correspondence between the basic elements of music (beat, note length) and CAD design elements (lines, shapes, colors). For example, strong beats in music can correspond to thick lines in CAD design, and weak beats can correspond to thin lines. Next, through computer vision technology, the visual features in music videos are extracted and the music rhythm is identified. According to the recognized rhythm, combined with the mapping rules, the relevant parameters and design elements in CAD design are automatically adjusted to realize the synchronous change with the music rhythm.

The realization of the system mainly includes the following steps: ⊖ Data acquisition: Obtaining music videos and initial CAD design files. ⊜ Visual feature extraction and music rhythm recognition: Using computer vision technology, visual features are extracted from music videos, and music rhythm is recognized. ⊛ Establishment of mapping relationship: According to the defined mapping rules, the corresponding relationship between music rhythm and CAD design elements is established. ④ CAD dynamic design: Based on the recognized music rhythm and mapping relationship, the parameters and design elements in CAD design are automatically adjusted to realize dynamic design. ⑤ Display and evaluation of results: Display the dynamic design results and evaluate the results according to the set evaluation criteria. In the running process, the system can receive users' input and adjustments in real-time to ensure the flexibility and interactivity of the design. At the same time, the system can automatically iterate and optimize the mapping relationship according to the needs so as to improve the accuracy and efficiency of the design.

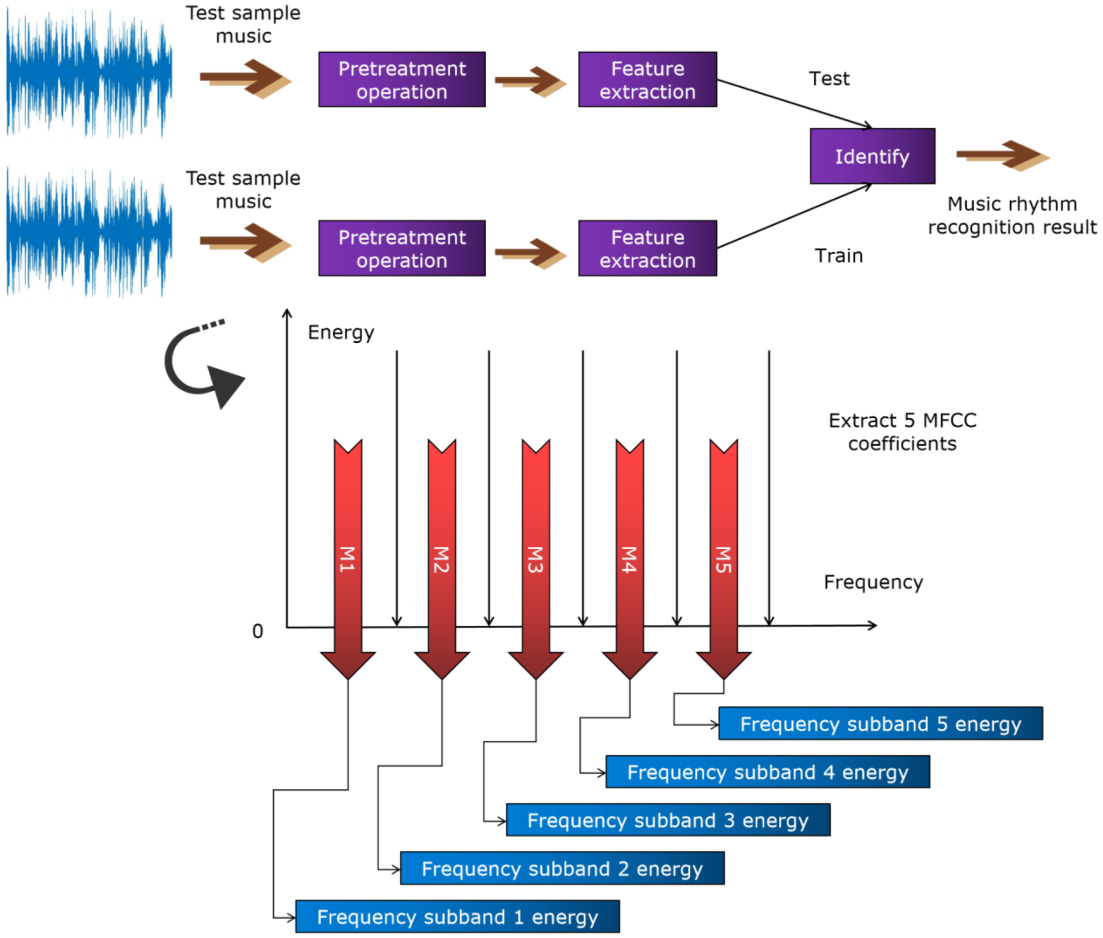

Figure 1: Linear MFCC coefficient extraction of audio short-time frames.

6 SIMULATION RESEARCH AND DESIGN EXAMPLE

6.1 Simulation Environment and Method

In order to verify the effect of the integration of music rhythm recognition and CAD dynamic design based on computer vision, this section builds a simulation environment for testing and verification. The simulation environment includes music and video data sets, CAD design software, a computer vision algorithm library, and so on. The experiment adopts the way of simulating music and designing tasks and simulates the process of music rhythm recognition and CAD dynamic design in real scenes. Figures 2, 3, and 4 are waveform diagrams of different simulated music signals.

In the simulation method, firstly, computer vision technology is used to preprocess and extract the features of the music video, and then the extracted features are identified and analyzed by the music rhythm recognition algorithm to get the music rhythm information. Next, the music rhythm information is associated with CAD design parameters, a parametric model is established, and CAD dynamic design is realized. Finally, the design results are simulated and evaluated.

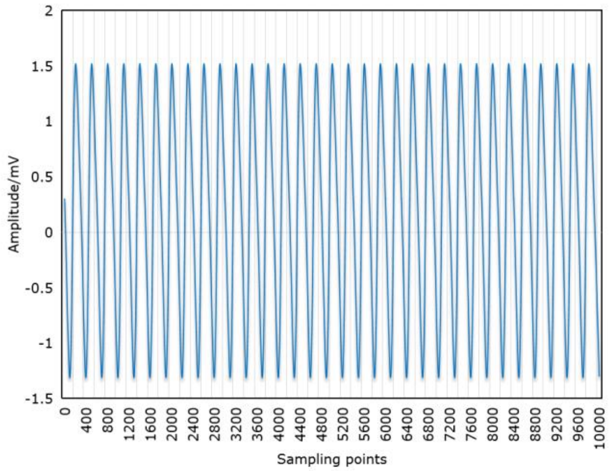

Figure 2: Music signal 1 waveform diagram.

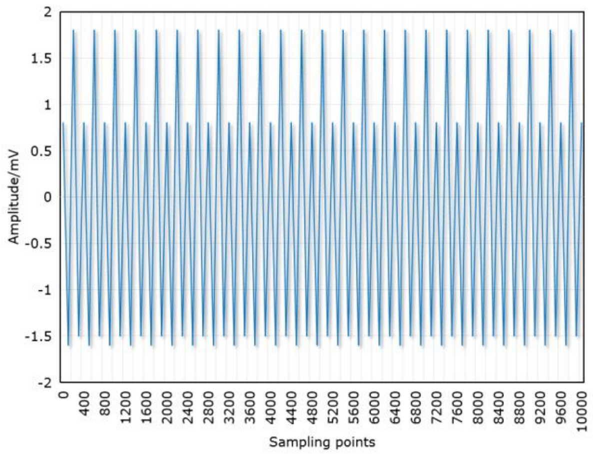

Figure 3: Music signal 2 waveform diagram.

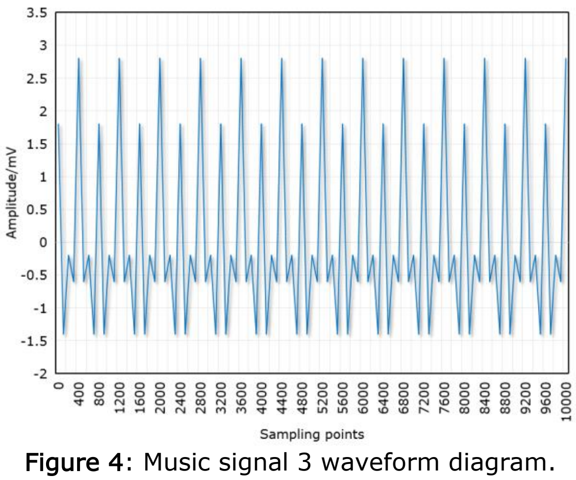

Figure 4: Music signal 3 waveform diagram.

6.2 Simulation Results and Analysis of Music Rhythm Recognition and CAD Dynamic Design

In the simulation of music rhythm recognition, this section uses a variety of different types of music videos to test, including pop music, classical music, and so on. Through computer vision technology to extract image features, combined with a rhythm recognition algorithm, the accurate recognition of music rhythm has been successfully realized.

In this section, in order to comprehensively evaluate the rhythm recognition performance of this method in different types of music videos, a variety of music video data sets are specially selected. This data set contains a variety of music types, such as pop music, classical music, folk music, and so on. Ensure that there are enough samples for each music type to accurately evaluate the performance of the model. This experiment mainly uses accuracy as an evaluation index to quantify the matching degree between the rhythm type predicted by the model and the real label. For each music type, the same preprocessing steps and CNN model are used for rhythm recognition. This ensures the consistency and fairness of the experiment. Figure 5, Figure 6, and Figure 7 show the accuracy of rhythm recognition in different types of music videos respectively.

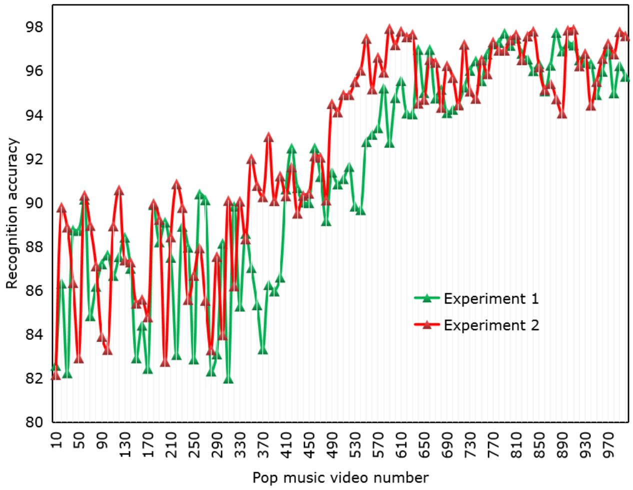

Figure 5: Accuracy of rhythm recognition in pop music videos.

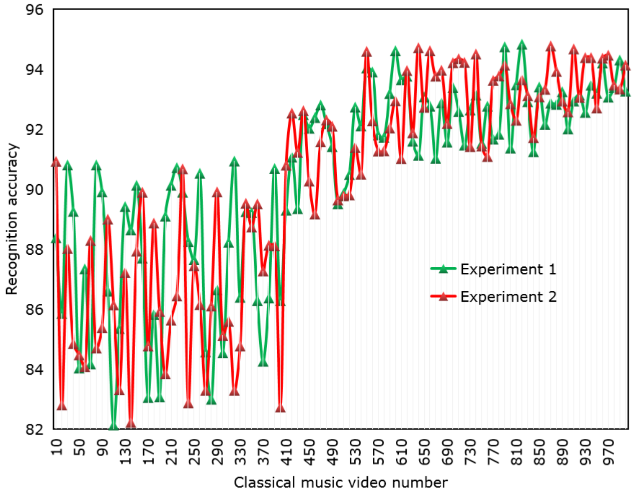

Figure 6: Accuracy of rhythm recognition in classical music videos.

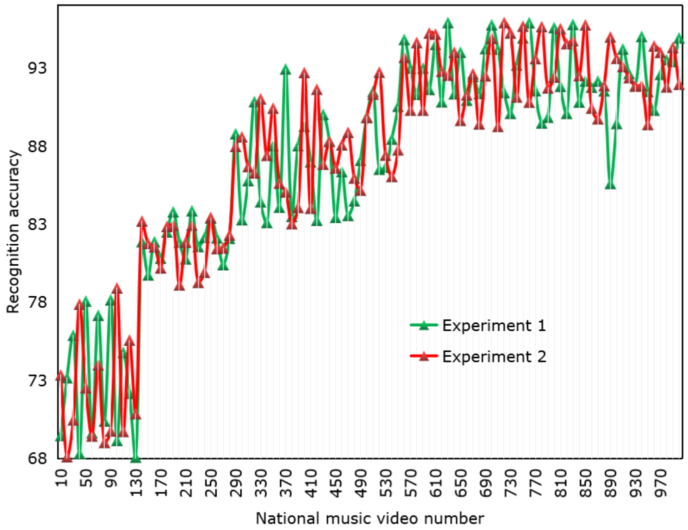

Figure 7: Accuracy of rhythm recognition in national music videos.

As can be seen from the results in Figure 5, the method in this paper has achieved high accuracy in pop music, which can basically reach above 97%. This is because the rhythm of pop music is relatively regular, and there are usually obvious percussion instruments, which makes the rhythm characteristics more obvious and easy to identify. Compared with pop music, the rhythm of classical music is more complicated and changeable, including more instruments and changes. Figure 6 shows that the accuracy of this method in classical music is slightly lower, fluctuating around 94%, but still maintaining a relatively high level. This shows that the method in this paper also has a certain recognition ability for complex music rhythms. National music usually has a unique rhythm and rhythm, which is different from pop music and classical music. As can be seen from Figure 7, the method in this paper has also achieved a certain accuracy in folk music, but it is slightly lower than the other two music types, about 92%. This is because the rhythm characteristics of national music are not consistent with the training data of the model, or special feature extraction methods are needed to better capture the rhythm characteristics of national music. On the whole, this method can realize high-precision rhythm recognition in different types of music videos. For complex music rhythms and changes, this method can also show good robustness and accuracy. This verifies the effectiveness and advantages of computer vision technology in music rhythm recognition.

In order to evaluate the time required for music rhythm recognition, this section uses a data set containing multiple music segments. These music pieces have different rhythm types and complexity to ensure the diversity and reliability of the experiment. The experiment was carried out on a computer equipped with a GPU to speed up the calculation process of the model, and the DL framework was used to build and train the CNN model. The code and algorithm in the experiment are optimized to ensure efficient calculation speed. Figure 8 shows the time required for music rhythm recognition.

It can be observed from Figure 8 that the time required for music rhythm recognition presents a relatively stable trend. Specifically, the rhythm recognition time of most music pieces is between 2 seconds and 5 seconds. This shows that the CNN model has high computational efficiency in dealing with music rhythm recognition tasks. This is mainly because by adjusting and optimizing the parameters of the model, the model can reduce unnecessary calculations while maintaining a high recognition accuracy, thus shortening the recognition time.

In the dynamic design simulation of CAD, according to the recognized music rhythm information, the CAD model is driven for dynamic design. By adjusting the design parameters and simulation technology, the CAD design that changes synchronously with the music rhythm is realized. The visual effects of different types of music are shown in Figure 9.

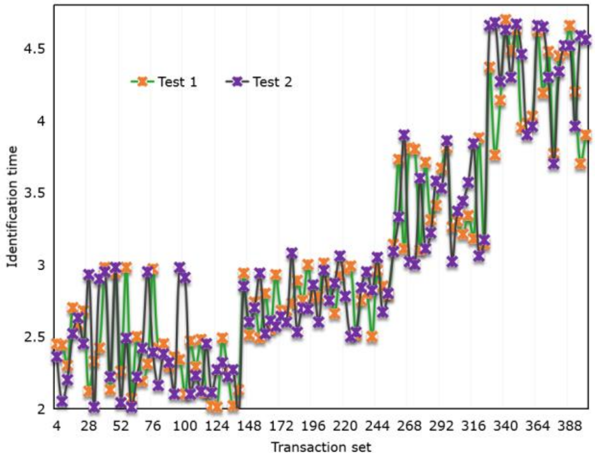

Figure 8: Time for music rhythm recognition.

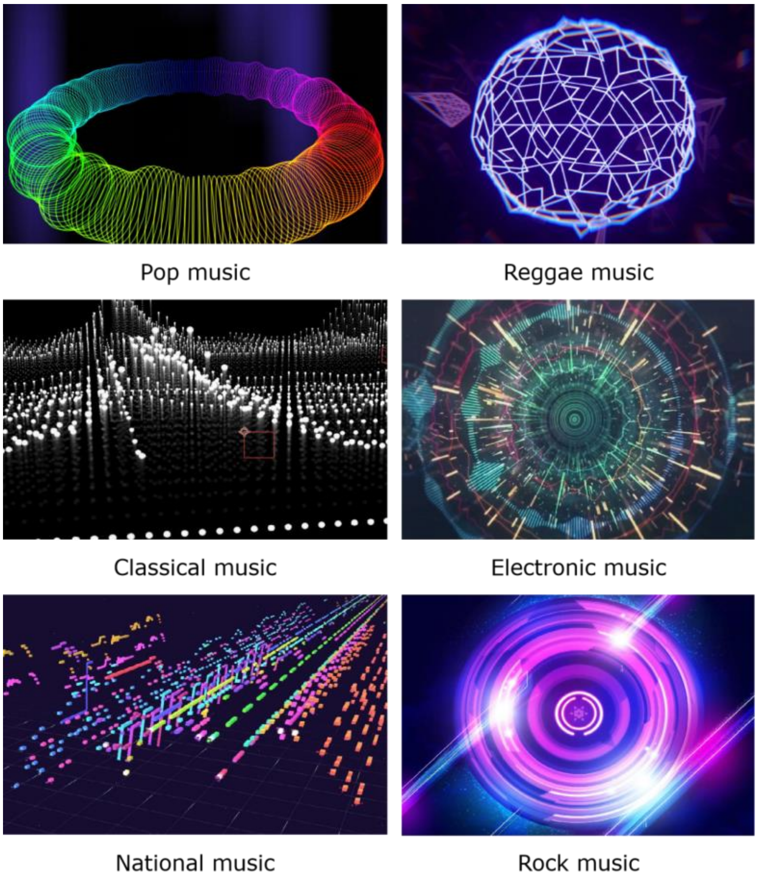

Figure 9: Visualization of different types of music.

The simulation results show that the method in this paper can reflect the change of music rhythm to CAD design in real-time, and realize the dynamic and interactive design. Designers can flexibly adjust the design parameters according to the change in music rhythm, and realize the close connection with music and creative expression.

The experimental results in this section show that the integration of music rhythm recognition and CAD dynamic design based on computer vision not only innovatively combines the technologies and methods of the two fields, but also provides a brand-new and cross-modal creation and expression way for designers and musicians, which promotes the deep integration of art and technology.

7 CONCLUSIONS

In this paper, computer vision technology is applied to music rhythm recognition and CAD dynamic design. Its main research includes: (1) In-depth study and selection of the existing computer vision technology, to find a method suitable for music rhythm recognition and CAD dynamic design; The music rhythm recognition model based on computer vision is constructed and verified by experiments. Study the core algorithm and technology of CAD dynamic design to realize a complete dynamic design process; The results of music rhythm recognition are combined with CAD dynamic design to verify its feasibility and practicability. The evaluation results show that the time required for music rhythm recognition using this method presents a relatively stable trend, and the rhythm recognition time of most music fragments is between 2 seconds and 5 seconds. At the same time, this system has achieved satisfactory performance in the accuracy of music rhythm recognition and CAD dynamic design. In practical application, the system can be widely used in music creation, stage design, architectural design, and other fields, providing artists and designers with brand-new creative tools and expressions. For example, musicians can use the system to create music videos that match the music rhythm, stage designers can use the system to realize the stage scenery that changes synchronously with the performance rhythm, and architects can integrate architectural design with the natural environment and humanistic rhythm through the system.

In a word, the CAD dynamic design method based on computer vision combines the advantages of computer vision technology and CAD design, provides designers with more intuitive, flexible, and efficient design tools and methods, and helps to improve the efficiency and quality of CAD design. Through in-depth research and application of computer vision technology, it is expected to make more innovations and breakthroughs in these fields.

8 ORCID

Fan Yang, https://orcid.org/0009-0008-8949-1315

REFERENCES

[1] Alfaro, C.-M.; Valero, M.-J.-J.: Exploiting the two-dimensional nature of agnostic music notation for neural optical music recognition, Applied Sciences, 11(8), 2021, 3621. https://doi.org/10.3390/app11083621

[2] Alhassan, H.; Al-Hassan, B.; Osei, M.: Comparative study of UCC and UEW B. Ed JHS Visual Art options as initial teacher preparation for the new creative arts and design curriculum for JHS, Journal Evaluation in Education (JEE), 4(4), 2023, 125-133. https://doi.org/10.37251/jee.v4i4.719

[3] Andrea, P.; Zahra, A.: Music note position recognition in optical music recognition using convolutional neural network, International Journal of Arts and Technology, 13(1), 2021, 45-60. https://doi.org/10.1504/IJART.2021.115764

[4] Ba, T.; Li, S.; Gao, Y.; Tan, D.: Research on the interactive design of electric vehicle interior based on voice sensing and visual imagery, International Journal of Vehicle Information and

Communication Systems, 8(1-2), 2023, 152-169. https://doi.org/10.1504/IJVICS.2023.131613
[5] Baro, A.; Riba, P.; Calvo, Z.-J.; Fornés, A.: From optical music recognition to handwritten music recognition: a baseline, Pattern Recognition Letters, 123(5), 2019, 1-8. https://doi.org/10.1016/j.patrec.2019.02.029
[6] Calvo, Z.-J.; Jr, J.-H.; Pacha, A.: Understanding optical music recognition, CM Computing Surveys (CSUR), 53(4), 2020, 1-35. https://doi.org/10.1145/3397499
[7] Cavdir, D.; Wang, G.: Borrowed Gestures: The body as an extension of the musical instrument, Computer Music Journal, 45(3), 2021, 58-80. https://doi.org/10.1162/comj_a_00617
[8] Garrido, M.-C.; Rios, V.-A.; Calvo, Z.-J.: A holistic approach for image-to-graph: application to optical music recognition, International Journal on Document Analysis and Recognition (IJDAR), 25(4), 2022, 293-303. https://doi.org/10.1007/s10032-022-00417-4
[9] Ge, M.; Tian, Y.; Ge, Y.: Optimization of computer-aided design system for music automatic classification based on feature analysis, Computer-Aided Design and Applications, 19(S3), 2021, 153-163. https://doi.org/10.14733/cadaps.2022.S3.153-163
[10] Maba, A.; Computer-aided music education and musical creativity, Journal of Human Sciences, 17(3), 2020, 822-830. https://doi.org/10.14687/jhs.v17i3.5908
[11] Paul, A.; Pramanik, R.; Malakar, S.; Sarkar, R.: An ensemble of deep transfer learning models for handwritten music symbol recognition, Neural Computing and Applications, 34(13), 2022, 10409-10427. https://doi.org/10.1007/s00521-021-06629-9
[12] Pei, Z.; Wang, Y.: Analysis of computer-aided teaching management system for music appreciation course based on network resources, Computer-Aided Design and Applications, 19(S1), 2021, 1-11. https://doi.org/10.14733/cadaps.2022.S1.1-11
[13] Ríos, V.-A.; Esplà, G.-M.; Rizo, D.; Ponce, L.-P.-J.; Iñesta, J.-M.: Applying automatic translation for optical music recognition's encoding step, Applied Sciences, 11(9), 2021, 3890. https://doi.org/10.3390/app11093890
[14] Rocamora, M.; Cancela, P.; Biscainho, L.-W.-P.: Information theory concepts applied to the analysis of rhythm in recorded music with recurrent rhythmic patterns, Journal of the Audio Engineering Society, 67(4), 2019, 160-173. https://doi.org/10.17743/jaes.2019.0003
[15] Tan, B.; Yang, F.: Dance movement design based on computer three-dimensional auxiliary system, Computer-Aided Design and Applications, 19(S3), 2021, 108-118. https://doi.org/10.14733/cadaps.2022.S3.108-118
[16] Wang, Q.; Zhou, L.; Chen, X.: Kernel density estimation and convolutional neural networks for the recognition of multi-font numbered musical notation, Electronics, 11(21), 2022, 3592. https://doi.org/10.3390/electronics11213592
[17] Wang, X.: Building a parallel corpus for English translation teaching based on computer-aided translation software, Computer-Aided Design and Applications, 18(S4), 2021, 175-185. https://doi.org/10.14733/cadaps.2021.S4.175-185
[18] Yang, D.; Goutam, A.; Ji, K.; Tsai, T.-J.: Large-scale multimodal piano music identification using marketplace fingerprinting, Algorithms, 15(5), 2022, 146. https://doi.org/10.3390/a15050146
[19] Yuan, Y.: Design and realization of computer aided music teaching system based on interactive mode, Computer-Aided Design and Applications, 18(S2), 2021, 92-101. https://doi.org/10.14733/cadaps.2021.S2.92-101