A Humble Opinion on the Computer Vision Core Technology
Current Situation and Prospect

Yi Feng
University of Kent, Canterbury
zhanghaipengzoe@163.com

Abstract. Recently, with the development and application of computer technology, research hotspots have gradually shifted to the intelligent extraction and analysis of rich visual resources contained in video data and digital images, so as to obtain effective information. Under such research background, computer vision technology has emerged and has been widely used in intelligent security inspection, manufacturing, human-computer interaction, image retrieval, and medical image analysis and other fields. However, there are still many problems in this technology, such as unstable image feature detection, low detection efficiency, and fuzzy description of semantic information. This paper has carried on the relevant discussion and analysis, and further explored its development trend.

1. Introduction
With the development of computer technology, researchers in many fields have tried to easily identify, locate, and even make relevant judgments and guesses through subtle changes in human beings through computer hardware and software and visual sensors, so that robot systems and computers can also simulate some intelligent visual functions such as image acquisition, analysis, processing, and learning.

Researchers have made constant attempts and efforts over the years, but still cannot achieve precision and flexibility compared to human visual levels. At present, this technology has become a hot spot of research and development.

2. Computer vision overview
The visual signals have more abundant information than physical information such as sound, which brings great complexity and difficulty to the description work. Complex information such as various human poses, scenes, object distribution, characters, facial features of people, etc., even if it is simple information, not only contains a large number of shapes, geometry, colors, shape features, etc., but also has greater subjectivity and dynamics. Therefore, the hot question of research is how to overcome various problems in the process of extracting image information [1].

At present, it mainly adopts a mechanism of inverse derivation based on known or hypothetical associations, and associates the output with the input, and guesses the shape, color, distribution, illumination, etc. of the physical object through the corresponding picture, for the mystery of the biological vision system has not been fully deciphered. Therefore, in this way of reversal, the field of technology more widely uses machine learning algorithms and statistical models. Compared with the
evolutionary process of bio-visual evolution of hundreds of millions of years, the statistical model and the learning mechanism, based on prior knowledge and based on the simplifications, constraints and assumptions of the image content are one-sided and rough. And the relationship between the input and output of the visual system can only be roughly described. After the increase of the number and type of samples in the machine learning mechanism, the stable mechanism is reduced, and the simple basic mathematical description performance is better than the complex detection algorithm. These are the results presented by the experiments using the reverse mechanism. And the mechanism will solve the difficulties and challenges in solving the problems, associated with this technology.

There are also forward derivation methods to solve some of these technical problems, but usually input artificial images, which are applied to book covers, barcode detection, and trademarks and so on. The information, manually controlled, contained in such images, maps all known one-to-one and controls the intrinsic link between its input and output information. This forward derivation mode is easier to solve for the technical problems of generalized natural images. Such technologies are mainly used in the practice of automatic book management systems and printing quality inspection systems to meet people's daily needs.

With the continuous development of technology in the field of computer vision, multi-granularity real-time extraction and tracking of human bones have been widely used in game development, film and television and other fields. The traditional optical three-dimensional information acquisition is mainly obtained by taking multiple images from multiple angles by the camera, with numerous requirements, complicated operation steps, long time to cost, limited promotion, and low acceptance.

Three elements of research and development are proposed in related, namely: reasonable constraints are introduced and the objectives of the mathematical calculation level are considered; the problem of data format concerns of input and output in vision is resolved, and the system function by rational design of the algorithm is realized; the hardware used needs to be considered whether it meets the algorithm requirements, and the counteraction to the description and the required algorithm.

3. System framework of computer vision

In the development of computer vision systems, it is necessary to make up for the semantic differences between the original image data and the knowledge it contains. It is determined by its inverse mechanism. Image content is reasonably assumed and simplified based on the satisfaction of specific application needs. For example, in the corresponding system, the assumptions used in various technologies are summarized, and the degree to which the summarized information can be used is close to the extent of actual application. Such a system can be divided into three levels of theoretical structure [2]. The relationship between these three levels is interconnected, mutually supportive, and has no absolute boundaries. This is also the system framework that is now widely used.

In the first layer of image data layer, the digital signal of the pixel level is mainly processed correspondingly, such as image compression, image acquisition and transmission, and noise reduction. The degree of requirements for simplifying the inspection environment and the implementation of the various specific algorithms applied subsequently are determined by the technology employed in this layer. For example, when the image feature is acquired, the difficulty of the object in the image segmentation scene is reduced by the now mature depth sensor, which significantly improves the efficiency of establishing the human body stick model. Directly through the three-dimensional feature positioning provided by the depth image, the human body posture recognition result can be effectively obtained, and the real-time performance and accuracy are excellent. It is extremely suitable for game product and interaction related application development.

4. Development history of computer vision

Computer vision technology has shown a rapid development in recent years. With the development of computer hardware and software technology and digital information processing technology in recent years, this technology has gradually become the theoretical framework of the system [3].
Computer vision application development must choose to use learning mechanisms and statistical features. The mainstream of computer vision development has become a comprehensive analysis and description of object information in images. At the same time, with the continuous development and advancement of Internet technology and computer software and hardware technology, visual algorithms previously restricted by computer level have been re-proposed, improved and optimized. In addition, other fields and disciplines that have been intertwined have also been greatly developed, such as two-dimensional image rendering technology of computer photography. Using corresponding algorithms and techniques, it is applied to digital cameras, capable of generating a dark background with high dynamic ratio but achieving a clear reproduction of digital photos.

5. Development trend of computer vision

Related research shows that, regardless of the description of the top-level semantic features or the underlying pixel-level feature description, this single image feature description mechanism limits the content of the image to a limited extent. At present, the main development trend in information description is to fuse multiple image features.

The combination of traditional local features and global image information structure forms a video event description factor that can combine local features with global features of images. This factor describes the relevant semantic relationships of complex multiplayer interactions in video in chronological order. Compared with a single large data, real-time performance has been greatly improved. The technology for identifying more complex image content that exists in the real world is also an important development and research trend. It is necessary to continuously develop corresponding recognition technologies, to be able to describe various features in real scenes, and to simplify, assume and constrain the image content. By using the hardware parallel processing mechanism, the local feature points are used for the reconstruction of the three-dimensional scene, which not only greatly improves the real-time performance of the algorithm, but also effectively improves the efficiency and accuracy of the scene recognition.

Various computer technologies, such as Internet technology, social media, and computer graphics technology, are continually integrated with the development of the Internet, and this is also a development trend that needs great attention. For example, the retrieval and learning of massive image content, the automatic annotation of image content, etc. can all be learned by means of the semantics of the computer, and the hole recognition degree of the image is also improved correspondingly with the semantic knowledge. The social relationship of characters in the image can effectively improve the accuracy of face recognition technology. This is suitable for the case where the relevant person around the target person is more likely to determine the recognition than himself.

6. Conclusion

At present, this technology has become the core technology and hotspot of research, and has become a new subject with broad application prospects. In some special fields, such as criminal reconnaissance, medical and health industries, higher requirements are put forward for this technology, such as: practicality, ease of use, effectiveness and stability. Therefore, this technology is facing huge development prospects as well as enormous challenges.

References

[1] Cao Boyu. Current Status and Prospects of Computer Vision Core Technology [J]. Research, 2016(12): 00199-00199.
[2] Yang Yan, Teng Guanghui and Wu Congling. Research Status of Computer Vision Technology in Discriminating Greenhouse Plant Deficiency [J]. Agriculture Engineering Technology: Greenhouse Horticulture, 2005(6): 40-41.
[3] Wu Lin. Overview of Computer Vision Navigation [J]. Computer Knowledge and Technology, 2008, 01(9): 143-149.
[4] Bian Jintian and Zhou Liyun. Path Recognition System for Mobile Robot Based on Computer
Vision [J]. Journal of Information Materials, 2017, 18(1): 49-51.

[5] Li Chuanjun, Li Xingcheng and Chu Shiyu. Application of Computer Vision Technology in Intelligent Parking Lots [J]. Computer Simulation, 2003, 20(5): 103-106.

[6] Liu Bing. Research and Analysis on Computer Vision Imaging [J]. Coal Technology, 2013(6): 177-179.