

Narration Framework of Chinese Ancient Fiction Images in the Digital Environment

Cong Jin, Shu-Wei Jin and Jin-An Liu

Abstract Narration of Chinese ancient fiction images has been concerned by many researchers. In the today of the digital technology rapid development, it will affect research of the image narration for Chinese ancient fiction. Based on the existing digital technologies, in this paper, an image narration framework in digital environment for Chinese ancient fiction is proposed. In the proposed framework, we analyze the possibility of using variety digital techniques for achieving the narration of Chinese ancient fiction images, whose implementation can provide support for the digital narration of Chinese ancient fiction images.

Keywords Chinese ancient fiction · Image narration · Image feature · Semantic description

1 Introduction

The digital engineering of Chinese ancient fiction started in the early 1980s, which has already achieved remarkable achievements. Currently, a large number of Chinese ancient fictions have been developed as the digital products with true meaning and have been successfully to the market [1–3], and the research about these digital products mainly includes discussing the current situation, development trend and researching countermeasure of the digitizing of Chinese ancient fiction, to introduce the achievements of the digitizing of Chinese ancient fictions, the used digital technologies and so on. However, there is few digital content research of Chinese

C. Jin (✉) · J.-A. Liu
School of Computer, Central China Normal University,
Wuhan 430079, Hubei, China
e-mail: jincong@mail.ccnu.edu.cn

S.-W. Jin
Département de Physique, École Normale Supérieure,
24, rue Lhomond, 75231 Paris Cedex 5, France

ancient fiction. In other words, the current main work of the digital engineering of Chinese ancient fictions is focused on the development of digital products. And the digital contents of Chinese ancient fiction after the product developed, in particular the digital images, were less researched.

The earliest appearing image in Chinese ancient fiction is in North Song Jia-You eight years (AD 1063) by Yu Jing-An written “Biography of ancient paragons”. After that, the images number in Chinese ancient fictions was gradually increasing, and which reached a very high level in the Ming and Qing dynasties [4]. The images in Chinese ancient fictions are a huge treasure house. It plays an important role for satisfying aesthetic needs of the readers, getting more visual information and enhancing cultural transmission capacity.

As an information carrier with rich semantics, the image includes richer information than the texts, which itself is easy to transcend cultural, ethnic and time barriers, and to transfer richer emotion and mood. Therefore, images have been more and more concerned and used, and they play an increasingly important role in many research and application areas. However, the digital engineering of Chinese ancient fictions is both a challenge and an opportunity for Chinese ancient fictions. People naturally hope that the Chinese ancient fictions are detailer studied and wider propagated by digital approaches, but which faces many problems.

More and more researchers believe that the digitizing of Chinese ancient fictions should not only reproduce their original copy, but rather the perfect combination of the modern technology and traditional content, and it should form a unified of tools and content [2–5]. Digitization of Chinese ancient fictions not only should be an adding value information base, but also should be an effective tool for academic research. So, it can provide the accurate statistical and semantic information with relating the content of Chinese ancient fictions and improve support function of researching Chinese ancient fictions. In existing the research of Chinese ancient fictions images, the digital approaches have not been fully utilized, which can not satisfy the current needs of the digital age. Therefore, the narration research [6] of Chinese ancient fiction images in the digital environment can find a new way for researching the image narration and may also provide an opportunity to enrich the current existing achievements.

2 Related Work

The narration is originally realized by language, and it is necessarily relates to image semantic content to achieve the narration. Therefore, from the digitizing, the premise of image narration is automatically to describe image semantic content, which relates to a standard description of an image metadata, needs description of image retrieval and content description of image semantic [7].

There are VBA, SVG, EXIP, MPEG-7 and so on [7, 8] in existing standards of image metadata. Generally, these standards are only suitable to describe the low

level features of an image, but it is usually very difficult to describe the semantic content of an image only using the low level features of an image.

Currently, the demand description method of the image retrieval may reflect the users' understanding for the images mainly by retrieval images. This description method can better reflect the deep content of images because these contents come directly from public users and the description of retrieval demand is relatively comprehensive.

Existing description methods of the image semantic contents can be used to classify image from image visual feature layer, image object space layer and image semantic concept layer respectively, which does not directly describe the semantic content of the image. For example, a description system of image semantic content based on natural language was proposed in [9]. However, this system can only describe relatively simple semantic content of the image, and its expression is not accurate. A description method of generating image semantic from the image annotation information was proposed in [10]. Drawback of this approach is that the description ability of image semantic content is limited, and the representation of image is also incomplete.

Through literatures retrieval, we found that there is not the digital research about the narration of Chinese ancient fiction images. In this paper, we will research the narration framework of Chinese ancient fiction images based on a variety of digital technology.

3 Semantic Description of Chinese Ancient Fiction Image

3.1 Feature Analysis of Chinese Ancient Fiction Image

Unlike general digital image, the images in Chinese ancient fiction were created by humans. Each painter has own painting style, each image contains creation of painter and shows feelings and thoughts of painter, and therefore there was a distinct personality creation feature. Furthermore, due to restriction of painting skills at that time, almost images in Chinese ancient fictions were binary only using lines for represent the image content. Therefore, an image in Chinese ancient fiction has not color feature. In other words, the image in Chinese ancient fiction only has texture and shape features. Its detail is shown in Fig. 1.

3.2 Semantic Description Model of Chinese Ancient Fiction Image

The standard description of image metadata, the requirement description of image retrieval and the content description of image semantic are fused to the semantic description model of the images in Chinese ancient fiction. Its detail is shown in Fig. 2.

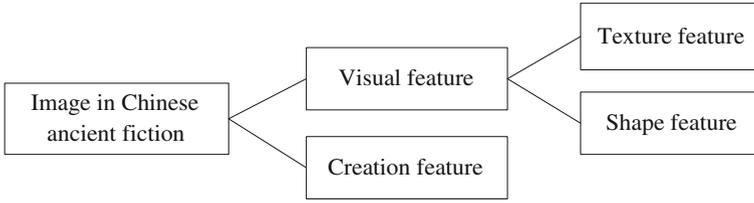


Fig. 1 Illustrating of the image features in Chinese ancient fiction

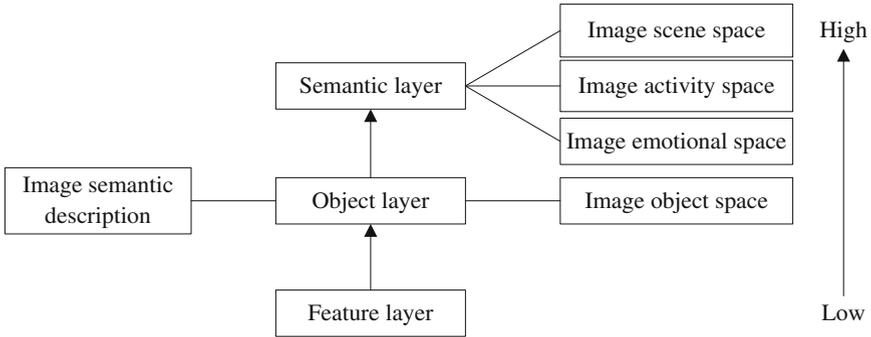


Fig. 2 Semantic description model of the images in Chinese ancient fiction

From the lowest “image feature layer” to the highest “image semantic layer”, the understanding of image contents is achieved also from the low level visual features to language description to describe the image content [11]. A detailed discussion of each layer is as follows.

(1) Feature layer

In Fig. 2, the lowest layer is feature layer in the semantic description model of Chinese ancient fiction image, including the creation features, shape features and texture features of an image. The creation features come from the people’s understanding for painting skills of an image creator, they belong to specialized knowledge and need to be put into the knowledge base. The latter two features belong to general concept, which are either pixel or set of pixels, and they can also be an abstract expression. Common characteristics of these features include point feature, line feature and regional feature. And their characteristics are as follows:

(1) Point feature

The position accuracy of point feature is very high, and its expression is very simple. But the number of the point feature is more, and the containing information is less.

(2) Regional feature

In regional feature, contains rich image information, itself number is relatively smaller. Its description is relatively complex, and position accuracy is poor.

(3) Line feature

The image information amount containing in line features is between point and regional features. Its computational cost belongs to moderate. Therefore, it is suitable for processing line image of Chinese ancient fictions. Extracting the shape and texture features can be automatically calculated by the computer. Common methods of extracting shape and texture features include edge detection, grayscale co-occurrence matrix, autocorrelation function of an image, Voronoi chess grid features, random field, Tamura texture features, auto-regression texture model, wavelet transform and so on.

Texture is a global feature, which describes the surface properties of the scene corresponding to the image or image regional. The texture features are not based on feature of the pixels, which needs the statistical calculation in a regional to contain many pixels. In the image matching, these regional features have some advantages, and therefore the local bias will not result in fail. Generally, texture feature is rotational invariance, and which has a strong noise resistance capability. However, there are drawbacks in texture features. For example, an obvious drawback is that the calculated texture may have larger deviations when changing the image resolutions. In addition, since texture is just a surface feature and does not fully reflect the essential attribute of the object, and therefore high level image content can not be obtained using only texture features.

Common shape feature extraction methods include boundary feature method, Fourier shape descriptor method, geometric parameter method, shape invariant moment method and so on. The description method of an image content based on shape features can more effectively describe the interest content of an image, but which has some problems. For example,

- (1) Currently, the image content description method based on shape features lacks more complete mathematical model as theoretical support, so sometimes the application results are not ideal.
- (2) When to exist the image distortion, sometimes the description results of image content are unreliable.
- (3) Many shape features only describe the local properties of the image content, and it requires more computing time and storage requirements for fully describing the content of an image.
- (4) The content information described by many shape features is not exactly the same with people's intuitive understanding. In other words, there is a difference between the similarity of feature space and the similarity of perceived by the human visual system and so on.

Therefore, in practical applications, it is very difficult only to use shape features for efficiently and accurately describing the content of the image, and requires the other features for better describing the image content.

(2) Object layer

Object is a target in image, such as people, animals, buildings or sky in image and so on. The part except the target is called the image scene. Image segmentation is a tool to obtain the targets of an image, which can divide an image into several targets with different features for further extracting information of the user interesting. There is spatial orientation relation, topological relation, and positional relation and so on between the targets. These relations to describe the image content are very important.

Spatial direction relation is mutually direction relation between multiple targets obtained by image segmentation, and these relations can be divided into the connection or adjacency relation, overlapping relation, inclusive relation and so on. Spatial topological relation describes the adjacent, relevance and inclusion relations between the points, lines and surfaces. The points, lines and surfaces can be used to describe connectivity, adjacency and regional between the targets. These topology relations are difficult to directly describe the spatial relation between the targets although adjacent but not link.

Spatial position information can be divided into two categories: the relative spatial position information and absolute spatial position information. The former relation emphasizes the relative case between the targets, such as above, down, left and right and so on. The latter relation emphasizes the distance and orientation between the targets. Obviously, the relative spatial positions can be obtained by absolute spatial positions, and the expression of the relative spatial position information is simpler.

Scene description is a general description of an image for other parts except the main targets, and its purpose is to avoid ambiguity problems of image semantics using the scene description method. Due to there is different understanding when different people to understand the same objective; it is inevitable that there is the ambiguity in the image semantics. In addition, since image has only two dimensional information in Chinese ancient fictions, there are differences with described the three dimensional world, which also led to difficulties to obtain the semantic only using the images. At present, the scenes in Chinese ancient fiction images can be divided into time scenes, such as spring, summer, autumn, winter, early, middle and late etc., and geographic scene, such as the ground, sky, indoor, outdoor, grasslands, deserts, oceans and so on. This knowledge can be put into the knowledge base.

Content description of object may complete part above work by using object recognition technology. It should be noted that the processing operation of object layer is based on the image segmentation. Since the target features acquired only using the image segmentation belongs incomplete features, only using these features can not carry on further operations, we also need some other descriptions to refer expertise or ontology of Chinese ancient fictions, and therefore these information also need to put into knowledge base.

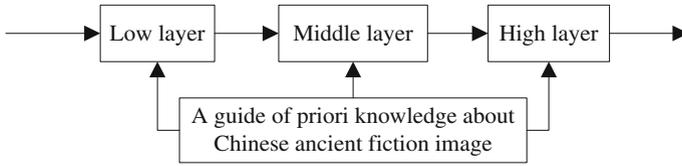


Fig. 3 Three levels of image understanding

(3) Semantic layer

In Fig. 2, semantic layer is at the highest level of image content description, which explains the image content or describes the image contents in natural language, and it also is called as image understanding or scene analysis [12]. Image understanding is consisted by two layers, the first layer is an image recovery scene, the second layer is to explain the image contents, namely high layer semantic of scene, and then they are matched with results of existing models using acquired knowledge.

Image understanding can be seen as a loop of some processing. The goal knowledge of the image content, all knowledge and the understanding experience of Chinese ancient fiction image may be stored in knowledge base. These processes of obtaining and storing knowledge are an actually learning process, and the process of image understanding can be seen as a process of matching and reasoning: After image processing, un-understood image is used to match targets within these images of the knowledge base. The background knowledge of these images of success matching within the knowledge base and all known knowledge and understanding knowledge about Chinese ancient fiction image can be used to understand those un-understood images for further inference and explanation.

The characteristics of image understanding are: information processing of several stages can bring multilayer represents of information, a correct understanding of an image needs guide of knowledge, and they can be described by the low, middle and high layers respectively. The detailed is seen in Fig. 3.

4 Narration Framework of Chinese Ancient Fiction Image

4.1 Obtain Topic and Analysis of Chinese Ancient Fiction Image

Images and text have them own to express topics, and to extract the image topic [13] is an essential task for understanding image. It is different that between to extract image topic by computer and the human eye. Image has narrative function, and the narration needs involve its topics, thus extracting image topics and further analysis these topics play an important role for image narration. It is essentially to

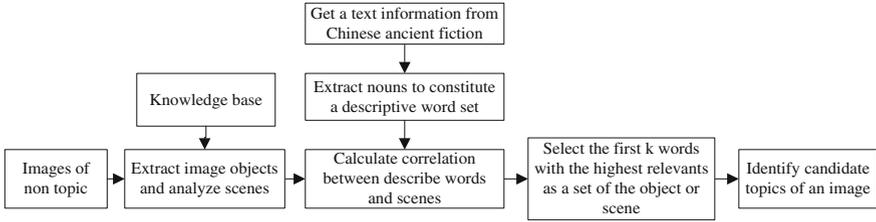


Fig. 4 Framework of extracting image topics

establish association between the image and text semantics, and which can also build a bridge for the future processing. Framework of extracting image topic is in Fig. 4.

In Fig. 4, the proposed framework of extracting image topics is mainly from the view of digital technology, and its result is differs from the topics directly given by the human eye. Since there is not only bias with the actual topics in the extracted image topics, and but also there is the ambiguous, synonyms, near-synonyms and so on, the extracted image topics by the framework of Fig. 4 can only be called candidate topics, it is necessary further to analyze the candidate topics.

The analysis of the candidate topics can be achieved by using natural language processing, mainly including the following two parts.

(1) Word sense disambiguation (WSD)

Besides there is a generally complex characteristic in the texts of Chinese ancient fictions, but also has its own characteristics of ancient Chinese vocabulary. For example, most of the words are ambiguous words [14] and so on. Therefore, WSD has a very high processing value. In particular description words of candidate topic, the meaning of polysemant is clear. WSD is a process to determine the meaning of ambiguous words for clearly describing them according to the particular candidate topics. The process of WSD is shown in Fig. 5.

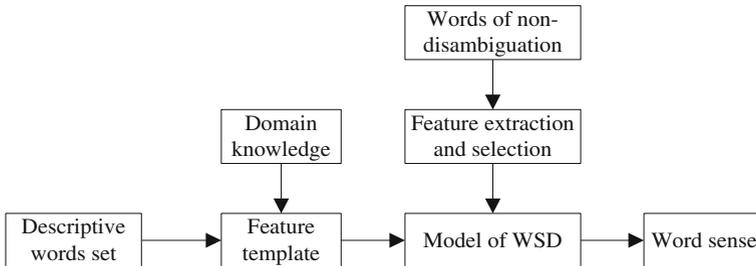


Fig. 5 The process of WSD

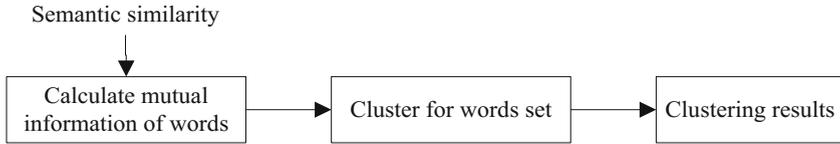


Fig. 6 The process of clustering

(2) Clustering of synonyms and near-synonyms

There are a lot of synonyms and near-synonyms in Chinese ancient fictions, the semantic similarity calculation between the Chinese ancient words plays an important role for clustering of synonyms and near-synonyms, and has a positive impact on the image narration. The semantic similarity of words reflects the correlation between words, and also reflects the semantic distance between words. Under the guidance of semantic distance between words, the clustering of synonyms and near-synonyms can be implemented. The process of clustering is shown in Fig. 6.

4.2 Time Model of Chinese Ancient Fiction Images

In [6], the author believes that the essence of image narration is time of space, which is that these images of spaced and decontextualized are put into a process of time for restoring or rebuilding their context. A lot of images in Chinese ancient fictions provide a good material for time of images, and digital research of narration of Chinese ancient fiction images also provides possible.

Narrative function of image necessarily involves a time series, because the narration is shown only according to time. Images have turned into a time slice of space media. For recovering narrative purpose, the movement of events must be reflected by many images, and these images must be incorporated into the process of time. Thus, all images of given a Chinese ancient fiction constitute a sequence of images according to the order of them appearing, so that we can time the spatial media, i.e., images. The narration model of Chinese ancient fiction images is shown in Fig. 7.

In Fig. 7, we add a time dimension for image, which allows that the image narration can reflect the movement of the events. Furthermore, in order to avoid unnecessary and contradictory text contents, after analyzing topics of each image, these topics continue to be processed for automatic summarization. The automatic summarization [15] of texts is a relatively mature technology, not repeats it here. In order to obtained summary sequence more smooth, the conflict resolution strategies [16] in artificial intelligence are also used in Fig. 7.

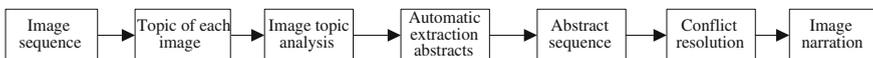


Fig. 7 The narration model of Chinese ancient fiction images

5 Conclusions

In this paper, the narration of Chinese ancient fiction images under digital environment is studied. According to the existing digital technology, we present a digital framework for the digital narration of Chinese ancient fiction images. Overall, the research of the narration of Chinese ancient fiction images with digital technologies is a complicated systems engineering, and which need integrate all aspects of various digital technologies. Although there are many difficulties, the proposed framework can play a positive role for automatically understanding the narration of Chinese ancient fiction images and its implementation may provide strong support for the digital research of Chinese ancient fiction images.

Acknowledgments This work was supported by national social science foundation of China (Grant No. 13BTQ050).

References

1. Chen, W.X., Wang, W.: Research of ancient Chinese fictions in digital times. *J. SE. Univ.* 7 (3), 118–121 (2005)
2. Wu, X.P.: Digitalization of ancient books and records and academic researches. *J. Guizhou Educ. Inst.* 23(6), 69–72 (2007)
3. Zheng, Y.X.: Influence and development direction of academic book digitization. *Manag. Rev. Soc. Sci.* (4), 81–88 (2006)
4. Li, F.L., Xun, S.: Research of Chinese ancient fiction image. *J. Ming-Qing Fict. Stud.* (4), 9–18 (2007)
5. Capital Library. Catalog of ancient fiction prints. Line Binding Bookstore (1996)
6. Long, D.Y.: Image narration: time of space. *Jiangxi Soc. Sci.* (9), 39–53 (2007)
7. Wang, X.G., Xu, L., Li, G.: Semantic description framework research on Dunhuang fresco digital images. *J. Libr. Sci. China* 40(209), 50–59 (2014)
8. Tzouvaras, V.: Multimedia annotation interoperability framework. <http://www.w3.org/2005/Incubator/mmsem/XGR-interoperability>
9. Li, Q., Shi, Z., Shi, Z.: Linguistic expression based image description framework and its application to image retrieval. In: Nachtgael, M., Van der Weken, D., Kerre, E.E., Philips, W. (eds.) *Soft Computing in Image Processing*, pp. 97–120. Springer, Berlin (2007)
10. Gupta, A., Mannem, P.: From image annotation to image description. In: Huang, T., Zeng, Z., Li, C., Leung, C.S. (eds.) *Neural Information Processing*, pp. 196–204. Springer, Berlin (2012)
11. Zhang, Y.J.: *Image Engineering*. Tsinghua University Press, Beijing (2013)
12. Huang, X.L.: *Image Retrieval Principles and Practice*. Communication University of China Press (2014)
13. Jin, C., Jin, S.W.: Automatic discovery approach of digital image topic. *Appl. Mech. Mater.* 598, 382–386 (2014)
14. Zhang, S.L.: *Ancient Chinese Knowledge Course*. Peking University Press (2002)
15. Zong, C.Q.: *Statistical Natural Language Processing*. Tsinghua University Press (2008)
16. Jin, C., Guo, J.L.: *Principles and Applications of Artificial Intelligence*. Tsinghua University Press (2009)



<http://www.springer.com/978-3-319-33623-7>

Artificial Intelligence Perspectives in Intelligent Systems
Proceedings of the 5th Computer Science On-line
Conference 2016 (CSOC2016), Vol 1

Silhavy, R.; Senkerik, R.; Oplatkova, Z.K.; Silhavy, P.;
Prokopova, Z. (Eds.)

2016, XV, 536 p. 158 illus., 72 illus. in color., Softcover
ISBN: 978-3-319-33623-7