

Scene Theme Extraction Based on Scene Lexicon

Fuping Yang^{1, a}, Yuke Bai^{2, b}

¹School of Computer Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

²School of Computer Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

^ayangfp@cqupt.edu.cn, ^bbbatworknow@outlook.com

Abstract: With the advent of the era of big data, the computing efficiency of computers has become higher and higher, and the artificial intelligence technology has been booming. These technologies have continuously developed to provide technical support for the text-to-scene conversion and brought new challenges at the same time. In the text-to-scene conversion system, there is a problem about that a text-to-scene conversion system can generate rich and varied scenes if the text input is strictly limited; if the strict restrictions on input are eliminated, the system can convert a vary kinds of scene. Therefore, if we want to a text-to-scene system can generate rapidly all kinds of scene, regardless of the length and the writing format of the articles, we need to know what scene theme of this article is. When we know the scene theme of this article, we can get the essential information when scenes are creating, such as the common entities under this scene, the attributes of common entities and spatial relationships of common entities. This paper brings out a notion of scene theme and expatiates a method to extract the scene theme from the article.

Keywords: text-to-scene conversion, scene theme, extraction of scene theme, scene lexicon.

1. Introduction

A word ‘text-to-scene’ was brought up by Bob Coyne in WordsEye system^[1,2,3]. A text-to-scene system can let ordinary users quickly create 3D scenes without having to learn special software or acquire artistic skills. The definition of text-to-scene is that the automatic visualization process from natural language to 3D scenes or animation. As show in Fig. 1, the text-to-scene conversion can be divided into two parts, one is the extraction of scene information from the text, the other is the generation of the scene. The first part uses natural language processing to extract the scene information such as entities, the attributes of entities, the relationships of entities including spatial relationships and non-spatial relationships. The scene information is all needed when generating scenes. The second part uses 3D model library to match the scene information and then generates scenes. The entities are the basic element of the scenes. The WordsEye system uses models to build 3D scenes in various themes. When scenes were built, system sign a theme randomly. Therefore, Words Eye system needs lots of models to meet all kinds of needs. Sylvain Dupuy from GREYC, a French research lab in digital sciences, created CarSim^[4,5,6] system based on WordsEye system. CarSim system uses car accidents as a specific theme to make the records of car accidents into 3D scenes. It can help police officers easily analyze the car accidents. Those text-to-scene systems show that building 3D scenes with a specific theme can make visualization process more easily. When the theme of scenes is identified, using entities in model library that match the theme, the relationships of entities and the environment or events that the words described can build a 3D scene. With the development of artificial intelligence a lot of text-to-scene system emerged, but there is no further study on scene theme. The paper firstly defines the scene theme then analyzes the features of scene theme and illustrates a way to extract the theme of the scene.

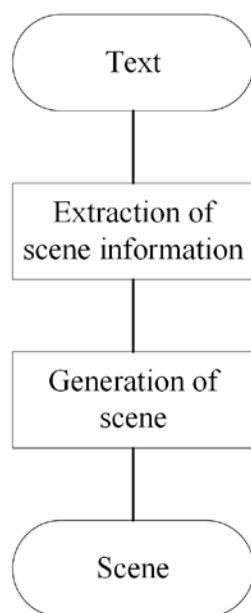


Fig. 1 The text-to-scene conversion

2. Scene Theme

To extract the scene theme of an article, we need know what scene theme is. Firstly, we collect articles with descriptions of the scene. Then we annotate the scene theme of each article and summarize the characteristics of scene theme. we will illustrate the characteristics of scene theme through the following aspects, the definition of the scene theme, the part-of-speech of the scene theme, the classification of the scene theme.

2.1 Definition of the scene theme

Text-to-scene is the process of the from text to scene automatically. The scene is built on what is described in the text. Scene should have a theme like the text. The connotation of the scene themes is similar to theme of text. And the scene theme is the summarization of the scene, the central topic of the scene. The definition of the scene theme is that a general description of a combination of many items or the connotation of an event in the reality world.

2.2 Part-of-speech of the scene theme

Part-of-speech means that a category to which a word is assigned in accordance with its syntactic functions. We analyze the part-of-speech of scene theme by collecting the theme words annotated in the corpus. According to the definition of the scene theme, the scene theme will be the name of item when the description of scene is about an item; The scene theme will be name of the event or summary of the event when the scene is about an event. It is obvious that no matter what scene theme is, name or summary, the part-of-speech of the scene theme will be noun. In some scenes which are describing an event, the event is dynamic at most of time. And in that cases, using a noun as scene theme cannot meet the requirement. We may use the verb-object phrases instead of noun. For example, when the scene describes a group of children flying the kites by the riverside. Using kite as the scene theme is inappropriate, flying kites will be the most suitable scene theme.

2.3 Classification of the scene theme

According to different division of scene, the scene theme can divide into different types such as static scene theme and dynamic scene theme, single scene theme and multiple scene theme.

It is common that the description of static scene is all about an item. The other items mentioned in the scene serve as a foil to the main item. So that the theme of static scene is the name of main item. Dynamic scenes often record about an event from beginning to end, and the theme of dynamic scenes may be the general description of events.

We can divide scenes into single scenes and multiple scenes on the basis of that the scene whether or not has other scenes in it. If the article describes a single scene, that means there is only a one scene,

the scene cannot be divided. If an article describes a scene with several sub scenes to make the article full, it is certain that this scene is multiple scene. And we name this scene main scene.

The theme of single scenes should be the item the scene described or the summary of events. The theme of multiple scenes depends on what scenes described. When scene depicts an item, and this item often involves a large concept. For example, when the article depicts the campus, the campus may divide into several parts such as playground, teaching building, garden and so on. The scene theme of this article should be campus, and the playground, teaching building, garden is the theme of each sub scene. When scene records an event, the sub scene may keep a record of the development of this event. The scene theme of main scene is general meaning of this event or the name of this event. The scene theme of each sub scene is summarization of what it describes.

3. Scene Theme Extraction

In this section the method of extract the scene theme is stated, including data sets, the building of scene lexicon, experiment design, experimental results and analysis in the following subsection.

3.1 Data Sets

In our experiment database is sourced from the Internet. We collect the compositions written by primary school students and encompassed different types of themes. These compositions mostly describe an event or an item so that all the articles can convert into scenes. The reasons why we use primary school students' composition are that they are easy to understand owing to the fewer texts, the simple sentence patterns, and low frequency of long sentences. With these features, the descriptions of the scene in the compositions and themes of scene can be analyzed easily.

The resulting database contains 1000 articles, all the articles have descriptions of the scene. We select 900 articles with 3 types of theme as training sets. Each type of theme has 300 articles, and they are campus, zoo and amusement park. The remaining 100 articles form the test sets. There are 20 articles for each theme listed above in the test sets. Besides, we add 40 articles with random themes different from those 3 themes in test sets to verify the accuracy of the proposed method.

3.2 Scene Lexicon

From the analysis of text-to-scene conversion, we know that the entity is the core of the scene. Even the scene theme may be the name of the entity. So that the scene lexicon can be built on the entities as follows. For each theme in the training set, firstly use the NLP words segmentation tool to segment the articles from the training sets. Then Tag Part-of-Speech of each word and filter the noun of each article. The articles transform into the bag of words. And make the statistics of frequencies of words in all the articles under one theme. The higher frequency of word, the more it can be explained that this entity is shown more often in scene under one theme. The process of building scene lexicon is shown in Fig. 2. And to show the results of building scene lexicon we choose first 10 term frequency of word under the scene theme is zoo as example, the results is shown in Table 1.

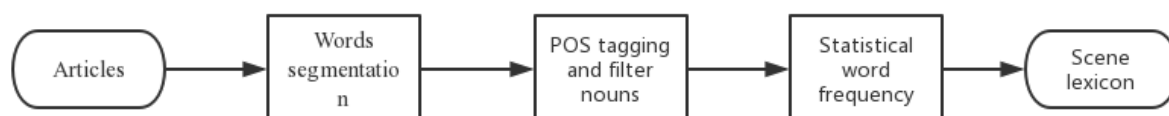


Fig. 2 Process of building scene lexicon

Table 1 Scene Lexicon

Scene Theme	Name of Entity	Word Frequency
Zoo	Zoo	0.11690836
	Animal	0.05247421
	Monkey	0.04260727
	Mother	0.03363731
	Tiger	0.02377037
	Father	0.02347137
	Peacock	0.0186874
	Elephant	0.01644491
	Lion	0.014800419
	Panda	0.00911945

3.3 Experiment design and results analysis

In our experiment all articles in the test set will be pretreated, including word segmentation, part-of-speech tagging, filtering out the noun. Then each article is as the same as the article in the training set. For each article all the words will match the words in the scene lexicon with one theme, after the matching finish, then move one theme to another until all the themes is matched with article. Next step is to calculate how many words are matched up with each scene theme. The more words are matched up in each article, the more possibility of scene theme is.

The results of experiment is that there are 60 articles' scene theme has been recognized, and 56 articles' scene theme are identified accurately. We use Precision, Recall and F-Measure as the evaluation criterion. And the $P=93.33\%$, $R=56\%$, $F=69.98\%$.

The results show that this method can extract the scene theme from the article, and some articles' scene theme cannot be recognized is due to the number of words in this article may be the same in different scene theme by matching up with the scene lexicon. To improve accuracy of this method, we add more article in the training sets. The results are show in Fig. 3. The accuracy is improved when the articles in training sets become more and more.

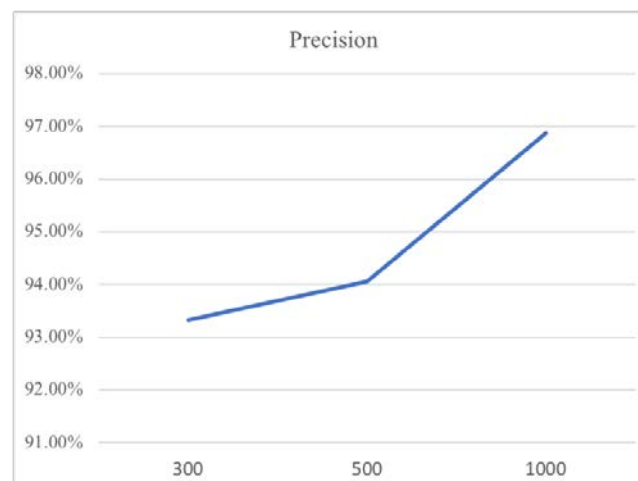


Fig. 3 The results of adding articles in training sets

4. Conclusions

In this paper, we proposed a notion of scene theme and illustrates a way to extract scene theme based on scene lexicon. The scene lexicon is the set of entities classified by the scene theme. By building the scene lexicon, we can recognize scene theme in articles. And the experiment shows that usability of extract the scene theme form articles by using scene lexicon. The accuracy of this method is improved when the training sets are increased. In future we attempt to add more scene theme in the scene lexicon. When the scene lexicon contains all kinds of scene theme, it will recognize any scene theme and make text-to-scene conversion much more easier.

5. References

- [1] Coyne B, Sproat R. WordsEye: An automatic text-to-scene conversion system[C]//Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques. United States: ACM, 2001: 487-496.
- [2] Coyne, Bob, Sproat, et al. WordsEye: an automatic text-to-scene conversion system[J]. 2001:487-496.
- [3] Coyne B, Sproat R, Hirschberg J. Spatial relations in text-to-scene conversion[C]//Proceedings of the Workshop on Computation Models of Spatial Language Interpretation at Spatial Cognition. Netherlands: CEUR-WS, 2010: 375-384.
- [4] Berglund A, Johansson R, Nugues P. A machine learning approach to extract temporal information from texts in Swedish and generate animated 3D scenes[C]//Proceedings of the 11th conference of the European Chapter of the Association for Computational Linguistics(EACL). United States: Association for Computational Linguistics, 2006: 385-392.
- [5] Johansson R, Berglund A, Danielsson M, et al. Automatic text-to-scene conversion in the traffic accident domain[C]//19th International Joint Conference on Artificial Intelligence. United States: International Joint Conference on Artificial Intelligence, 2005: 1073-1078.
- [6] Johansson R, Williams D, Berglund A, et al. Carsim: A system to visualize written road accident reports as animated 3D scenes[C]//Proceedings of the Second Workshop on Text Meaning and Interpretation. Association for Computational Linguistics, 2004: 57-64.
- [7] Dupuy S, Egges A, Legender V, et al. Generating a 3D simulation of a car accident from a written description in natural language: The CarSim system[C]//Proceedings of the Workshop on Temporal and Spatial Information Processing. Association for Computation Linguistics, 2001: 1-8.
- [8] O'Kane M, Carthy J, Bertolotto M. Text-to-scene conversion for accident visualization[C]// 2004:63.
- [9] Liu Z Q, Leung K M. Script Visualization (ScriptVis): a smart system that makes writing fun[J]. Soft Computing, 2006, 10(1):34-40.
- [10] Winograd T. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language.[J]. Computational Linguistics, 1971:464.
- [11] Simmons R F. The clowns microworld[C]// The Workshop on Theoretical Issues in Natural Language Processing. Association for Computational Linguistics, 1975:17-19.