

## Interactive book cover design based on language resources

Ki-Suk Lee\*,  
Virtual Reality Research  
Team,  
Creative Contents  
Research Laboratory,  
ETRI, Korea  
E-mail : mvr\_lks@etri.re.kr

Su-Ran Park,  
Virtual Reality Research  
Team,  
Creative Contents Research  
Laboratory,  
ETRI, Korea  
suepark@etri.re.kr

Hyun-Bin Kim,  
Virtual Reality  
Research Team,  
Creative Contents  
Research Laboratory,  
ETRI, Korea  
hbkim@etri.re.kr

Young-Kuk Kim,  
Computer Science  
& Engineering  
Department,  
Chungnam National  
University, Korea  
ykim@cnu.ac.kr

### 1. Introduction

For a long time, image has been used to communicate with each other because it is useful to describing a situation with background information of story. In this paper, we performed studies to represent summarizing storybook in a single image such as a movie poster or a book cover. The motivation of undertaking this study is to convert a text-based content to a multimedia content for children who are more familiar to image than text. It improves children's understanding and interesting of the story. The most of the previous researches were used to specify the layout of a scene from analyzing spatial information of a structured input text but we are based on an unstructured text of a commercial story book [1], [2], [3]. The main purpose of this research is to develop an easy-to-use interactive authoring tool for image generation by analyzing words in a story book using a natural language processing.

### 2. System Overview

The interactive book cover design system is an editor based on natural language processing and Windows-based GUI. This system is to provide a generated image from a book story and image-authoring environment for publishers and public users. This system comprises three major modules: 1) Natural Language Processing module; 2) Graphical User Interface (GUI) module; 3) Resource Database module.

The workflow of the system is as follows: the user loads a text of the story into the system. Then the Natural Language Processing module analyzes the text and sends the Resource Viewer a result of candidate lists from the Resource Database module. The Resource Database module stores Visual resources that are gathered from pre-built or web query. Stored data is integrated with indexing data to find visual resource from the text and the Knowledge DB. After that, the user composes a scene by selecting candidates using the Graphical User Interface which supports interactive editing without scripting or programming. The user modifies a pose of the character using a predefined pose data or the Pose Editor based on inverse kinematics.

Finally, when the user finishes scene editing, the system stores the output image and the user selection. Stored data was made up of an image layout and a correlation between visual resources and words. It will reuse data as an image template and change the ranking of candidates in the next time. As system utilization is increased, growing of image template will enable a more efficient job.

We show the relationship and workflow of each module in figure 1.

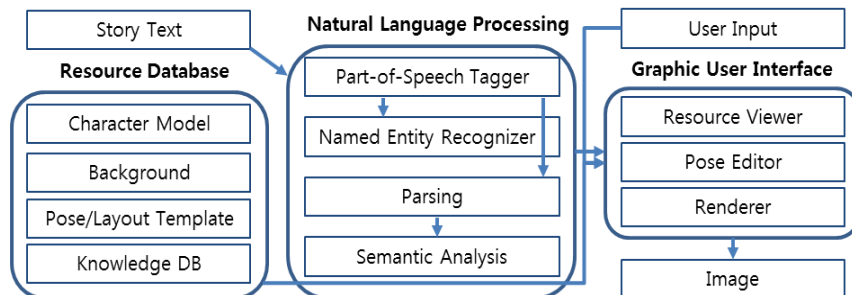


Figure 1. System Diagram and Workflow

### 3. Image Generation from an Unstructured Text

Developing an authoring tool using a natural language processing has many problems for working in a fully automatic manner because there are many ambiguity problems in an unstructured text [4], [5]. It is also same problems in understanding of human communication. In addition, when we imagine a scene from a story text, there are always many different opinions between users. It is caused by many lack of information to visualize image in a story text itself. Therefore, our developed system employs interactive method for solving problems in extracting information from a story book. For this purpose, the natural language processing module makes candidates for assigning parts of an image by analyzing text information. Candidates are each parts of the scene such as character, background, pose of the character and layout template. For finding character and background, we select all nouns as candidate using a part-of-speech tagger and then filter out by named entity recognizer. Candidates are sorted by the number of occurrences of each word contained in a text. And the others of the candidates for setting up the scene are determined by measuring the score of finding good resources for the text by the bag of words method. The bag-of-words Technique is widely used in natural language processing and based on the histogram of the frequency of words.

Each time the user changes the selection of the candidate, system modify a recommendation list. For example, when user selects a candidate of background, system updates the candidates in scope which related about selected background. For better result, we make various rule and knowledge of the image layout to fill the lack of visualize information. Our study use words in over 400 books to classify the visual resources and to build the knowledge base. Figure 2 shows the interface of the system and the final output image for the book cover.



Figure 2. Developed System Interface and Output Result

### 4. Discussion and Future Work

We present authoring system easily to generate image for the book cover. User can make an image from each parts of the scene intuitively by interactive selection and drag-and-drop method. Our system advantage is to find the visual resources by step-by-step user selection from initial loading text. Our study is a meaning for providing the method to produce content in a variety ways, to be able to make by user who is no talent in drawing. We expect many benefits in aspects of retrieval and rapid prototype, reusable resources trough connecting semantic word and visual resource. In order to generalize our technique, we need a lot of visual resources and shall expect to gradually resolve through development of big data and recognition technology. We will expand our work to collaborate creating an image with a remote user via Network.

### 5. References

- [1] B. Coyne and R. Sproat, "WordsEye: An automatic text-to-scene conversion system", In Proc. of SIGGRAPH, pp. 487-496, 2001.
- [2] X. Zhu, AB. Goldberg, M. Eldawy, et al, "A text-to-picture synthesis system for augmenting communication", Association for the Advancement of Artificial Intelligence, Vol. 7, 2007.
- [3] SR Park, KS Lee, SW Ryu and HB Kim, "User-friendly 3D Illustration-generation System for Fairytales", Proceedings of the International Conference on Convergence Content, 2013
- [4] A. Deshkar, A. Dhole and P. Vyas, "Natural language words analysis for affective scene generation from written text using artificial neural network. ", Journal of Engineering and Technology, pp. 1-5, 2012.
- [5] HS Kim and JC Park, "Real Time Synthesis of Multimedia Tales in Korean with Combinatory Categorical Grammar", Proceedings of the National Conference on Korean Information Processing, 2001
- [6] This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research & Development Program 2014. [R2013030002]