

Visual Dynamics Model for 3D Text Visualization

Sooyeon Lim

Department of Game

Dongyang University, Dongducheon, 11307, South Korea

ABSTRACT

Text has evolved along with the history of art as a means of communicating human intentions and emotions. In addition, text visualization artworks have been combined with the social form and contents of new media to produce social messages and related meanings. Recently, in text visualization artworks combined with digital media, communication forms with viewers are changing instantly and interactively, and viewers are actively participating in creating artworks by direct engagement. Interactive text visualization with additional viewer's interaction, generates external dynamics from text shapes and internal dynamics from embedded meanings of text. The purpose of this study is to propose a visual dynamics model to express the dynamics of text and to implement a text visualization system based on the model. It uses the deconstruction of the imaged text to create an interactive text visualization system that reacts to the gestures of the viewer in real time. Visual Transformation synchronized with the intentions of the viewer prevent the text from remaining in the interpretation of language symbols and extend the various meanings of the text. The visualized text in various forms shows visual dynamics that interpret the meaning according to the cultural background of the viewer.

Key words: Text Visualization, Visual Dynamics Model, Interactive Text.

1. INTRODUCTION

As new media based on digital technology develops, the use and utilization of text are rapidly changing. The text in the new media age adds cognitive functions to the essential functions that convey messages and generate meaning, and extends the usability and versatility as a digital contents material. Text is a creative symbol created to visually show everything that people think and explore. Text has evolved along with the history of art as a means of communicating human intentions and emotions. Since the 1970s, the texts of works of art have combined social forms and new media content, conveying social messages in capitalist society. Text visualization combined with digital media transforms the form of communication with viewers into an immediate and interactive form, enabling viewers to actively participate in the creation of artworks.

Interactive text visualization is the technique by which the viewer interacts with text with elements of movement for the purpose of communication. It is the addition of interaction with the viewer to the technique of transforming the color, size, position, etc. of text dynamically with time. In an interactive multimedia space, text is shaped by dynamic, interactive motion and instant communication. In other words, interactive text visualization has the characteristic of conveying information in real time with esthetics and creativity, and

exchanging emotion with the viewer. Therefore, with the popularization of multimedia and the formation of a culture in which video, sound, and text are used in general, attention is increasingly focused on interactive text visualization, which is a communication environment, and many artworks are being created. However, existing interactive text visualizations do not reflect the active interaction of the viewer because they define a few motion algorithms basically and allow only a small amount of editing within the range. And it is language-specific and does not support multiple languages.

This study proposes a theoretical model that produces language-independent and user-friendly interaction with user intention and human meaning, and an interactive text visualization system based on this model. For this purpose, an algorithm has been proposed that recognizes text as an image and preprocesses it and interacts with the viewer using motion recognition technology. The proposed system is an interactive text visualization that moves according to the viewer's gesture by using the skeleton information of the viewer tracked by the depth camera to synchronize with the text. This will enable not only the function of information transmission but also the higher level communication that can express emotional expression.

The remainder of this study is organized as follows. In Sect. 2, we review some researches related with the proposed system. In Sect. 3, the proposed visual dynamics model is presented in detail. We describe the process of text visualization authoring tool based on proposed model in Sect. 4 and conclude this study in Sect. 5.

* Corresponding author, Email: sylim@dyu.ac.kr
Manuscript received Sep. 27, 2018; revised Dec. 21, 2018;
accepted Dec. 21, 2018

2. RELATED WORKS

Beyond the oral period, the text that has crossed the character era and the digital era has been transformed by the evolution of media and technology. In media arts using digital media, various methods were used to image the text containing the intention of the artist or viewer and to convey the message, emotion or meaning contained in the text.

An example of an artwork that induced mental interaction without physical interaction with the viewer is Jenny Holzer's 'For the Guggenheim' [1]. She communicated with the public by publicizing texts containing controversial messages from politics, advertising, and popular culture in the real world as works of art in the streets of the city rather than in the museum. This artwork is a media façade to commemorate the 50th anniversary of the opening of the Guggenheim museum in New York. People who viewed this artwork felt and responded to the emotions associated with it through reading, discussing, and communicating text that was projected in public places.

One example of a viewer's intentional manipulation of reading and viewing text is Romy Achituv and Camille Utterback's 'Text Rain' and 'Written Forms' [2], [3]. These are artworks in which the text responds to the movement of the viewers who have been traced by the camera. In this work, viewers interact with each other by moving their bodies according to the emotions they feel in real time.

There is an artwork that uses the interaction of the user to change the attributes of the text from visual to tactile. R. Yoshioka developed a '3D Kanji' system in which texts are expressed emotionally in 3D space [4]. When a user enters a 2D text and his or her current emotion, a 3D modeling source of the text according to a variety of predefined emotion types is provided and 3D printed. Thus, this system satisfies the creative desire of users who want to model their ideas and engrave their artwork.

When a person expresses his/her emotions in text and uses it to communicate with other people, it is often transmitted differently from his / her intention. This is because the feelings and emotions about text can vary from person to person. If the feelings and emotions in the text are different, the meanings of the text also change. Therefore, it is necessary to study emotional models that objectify and quantify emotional words.

The most influential research on emotional models that visualize human emotions is Russell's emotional model [5], which defines emotional vocabulary as a two-dimensional graph. In this study, Russell describes a two-dimensional model based on Arousal and Valence, which is not only a form of analysis of the relationship between vocabulary and a list of emotional vocabulary but also a basis for quantifying the emotional state respectively. Beginning with Russell's research, studies of expressing emotion information in multidimensional graphs in various disciplines (medicine, psychoanalysis, engineering, etc.) studying human emotion have begun to become popular. Many recent studies have been proceeding by expanding the two-dimensional emotion model into a three-dimensional space.

The PAD emotional state model [6] is a psychological emotion model developed by Albert Mehrabian and James A. Russell. PAD expresses the sensibility to express in 3D space

with three axes (pleasure, arousal, dominance). Lovheim cube of emotion model [7] expresses emotional information by combining three kinds of hormones (dopamine, noradrenaline, serotonin) and eight basic emotions as a 3D emotion model. These 3D emotional models are not a new model as in most of the existing studies, but rather represent a multidimensional representation of the relationship between emotional words based on Russell's emotional model.

Therefore, these studies are somewhat lacking in expressing the polymorphism of modern multimedia texts and the various emotions embedded in them. For this reason, it is necessary to define an emotion model optimized for the meaning representation of multimedia texts [8].

3. VISUAL DYNAMICS MODEL

In this study, we propose an emotional model that can express the meanings and their relation to various forms of text. Visual dynamics of text means that various meanings of text are generated with the change of text form and various interpretation of the viewer occurs. In order for the viewer to interpret the text in various ways, his feelings must be involved in the interpretation process. The proposed Visual Dynamics Model transforms and reconstructs the expression form of the text for the production of various meanings of the text.

Fig. 1 shows the proposed visual dynamics model. In order to create the visual dynamics of the text, three elements (the viewer's intention embedded in text, the transformation of the text and the activity of the viewer to cause the transformation of the text) are essential. In this model, the visual dynamics of the text are expressed in space with a three-dimensional axis (X: Intention, Y: Activity, Z: Distortion). The X-axis represents the affinity of the intent of the viewer associated with the text, and the Y-axis represents the degree of movement of the viewer. And the Z axis represents the degree of distortion of the text. Because the interpretation is different for each viewer even for the same text, various text distortions are required for the production of visual dynamics. So, the visual dynamics are expressed more and more as the produced text meanings spread out in this model.

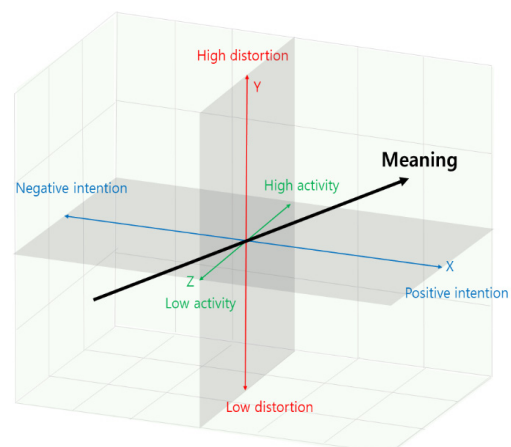


Fig. 1. Visual dynamics model of text

The intent of the viewer embedded in the text depends on the cultural background of the viewer and is involved in the production of new meaning. In order for the text to be interpreted differently according to its form, it must be combined with the feelings of the interpreter. Other interpretations for the same text are possible because the physical and psychological background of the viewer is different. People are not free from their social conditions and life experiences. The psychological and subjective experience of a viewer is closely related to the interpretation of the text and is revealed to the outside by emotion or action.

In this study, we applied 28 emotional words classified by Russell to the proposed model and used it as an element for analyzing the meaning of text. Fig. 2 shows how 28 emotional words are arranged in our model.

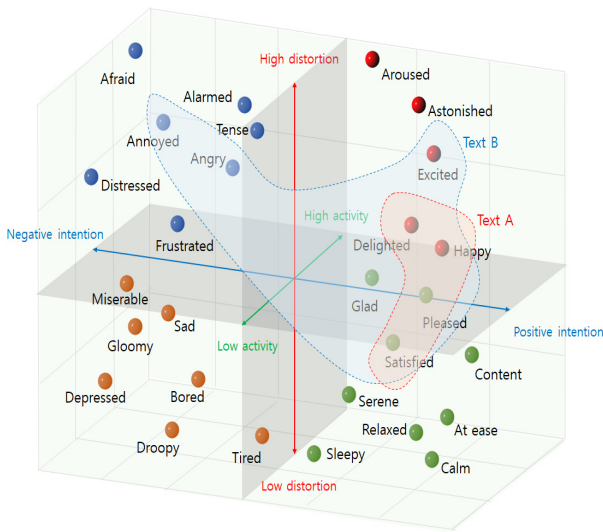


Fig. 2. 28 emotional words in visual dynamics model

In the 3D graph in Fig. 2, the area connecting the emotional words that the text B can express is wider than the area connecting the emotional words that the text A can express. This means that the greater the distance between the meanings of the texts, the higher the visual dynamics of the text. Distorted text shapes created by the viewer's interaction produce completely different meanings related to the viewer's background, current emotional state, and so on. In other words, distorted shapes created from a text are the result of visualizing the viewer's intentional interactions and various changes in emotions.

The meaning m_i generated by the text in the proposed model is defined as a vector in space R^3 . It consists of three components I, A, D .

$$\vec{m}_i = (I_i, A_i, D_i)$$

The degree of production of visual dynamics (VD) is defined as the variance of the meanings produced by the transformed text. Where m_{av} is the mean value of the meaning vectors produced as follows.

$$VD = \sqrt{\frac{1}{n} \sum_{i=1}^n (\vec{m}_i - \vec{m}_{av})^2}$$

$$= \sqrt{\frac{1}{n} \sum_{i=1}^n ((I_i - I_{av})^2 + (A_i - A_{av})^2 + (D_i - D_{av})^2)}$$

$$m_{av} = \frac{1}{n} \sum_{i=1}^n \vec{m}_i = \frac{1}{n} \sqrt{\sum_{i=1}^n (I_i^2 + A_i^2 + D_i^2)}$$

Visual dynamics refers to the state in which the external meaning of the text itself and the internal meaning derived from it are constantly produced. In this study, we divide visual dynamics into external dynamics and internal dynamics. The text has readability, so the external meaning of the text itself is continuously evoked to the viewer and the shape is transformed. This is external dynamics. In addition, the distorted form of the text produces the internal meaning associated with the viewer's personal background, along with the deconstruction of readability. This is the inner dynamics. By the viewer's interaction, both dynamics form a cyclic loop and produce various meanings related to text. Fig. 3 outlines the production process of visual dynamics. This figure shows that the distorted form of the text can be linked and reproduced with various meanings related to the background of the viewer interpreting it.

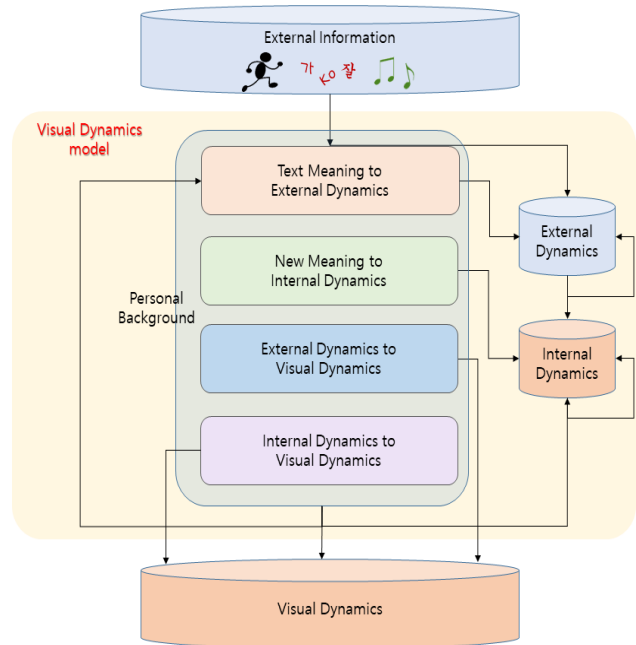


Fig. 3. The production process of visual dynamics of text

Produced new meanings are repeated to connect to other meanings through interaction with the viewer and create an internal space composed of connected meanings. The viewer's interaction also creates an external space through external transformations of the immediate text. In this way, the visual

formality of the text newly given through the overlapping and deconstruction of the text image provides the viewer with spatial depth and visual pleasure of the text image, thereby inducing viewer to actively participate in the artwork and producing various related meanings.

4. 3D TEXT VISUALIZATION ON PROPOSED MODEL

This study aims to develop a text visualization authoring tool that interacts instantly by user gestures based on the proposed visual dynamics model. To this end, we made an Interaction model that induces rich visual representation and meaning reproduction of text in 3D environment with interaction of the viewer.

4.1 Interaction model

The proposed 3D text visualization draws various meanings embedded in the text using 3D interaction with the viewer. The 3D interaction with the viewer is a three-step combination of access, communicate and create. Fig. 4 is an interaction model for producing visual dynamics [9]. The first step, access is the process of accessing text and recognizing it as a visual symbol. Attributes such as the font, color, and size of the text have a great influence on viewer recognition process. The second step, communicate, is the most important step in the production of visual dynamics as a process of understanding critically and interacting with the viewer's perceived text. The viewer looks at the forms of the text and makes associations according to his various cultural backgrounds and understands the meanings inherent in the text. And the viewer learns its meanings through repetitive interactions. The final step, create, is the process of producing new shapes and meanings of text as a result of viewer interaction. The viewer interacts more positively with randomness, contingency, and improvisation while viewing the transformations of the text, which is his or her creation.

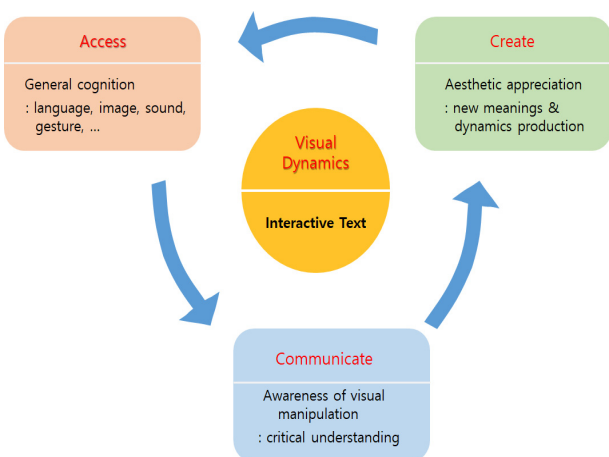


Fig. 4. Interaction model for visual dynamics production

4.2 3D text visualization

The 3D text visualization tool proposed in this study has features that the viewer inputs the desired text and interacts with the text in real time. In the proposed tool, text is imaged

and deconstructed in order to visualize input text as independent rather than dependent on a specific language.

Fig. 5 shows the procedure of the proposed 3D text visualization tool. It consists of three processes (Text deconstruction, 3D interaction, 3D text transformation) to produce visual dynamic of text. The first step, text deconstruction is a process in which the imaged text is performed by the thinning and contouring process and then deconstructed based on the inflection point to extract interaction points. These interaction points are used for communication with the viewer. The second step, 3D interaction, extracts the skeleton points of the viewer's gesture tracked by Kinect, the depth camera. These skeleton points are connected using the nearest matching algorithm with the interaction points obtained in the previous destruction process. In the final step, 3D text transformation, the interaction points of the text and the skeleton points of the viewer are matched n to n , and the distorted text is output by reconstructing the shape of the text. The various forms of the newly produced text generate repetitive external dynamics and internal dynamics and produce related new meanings.

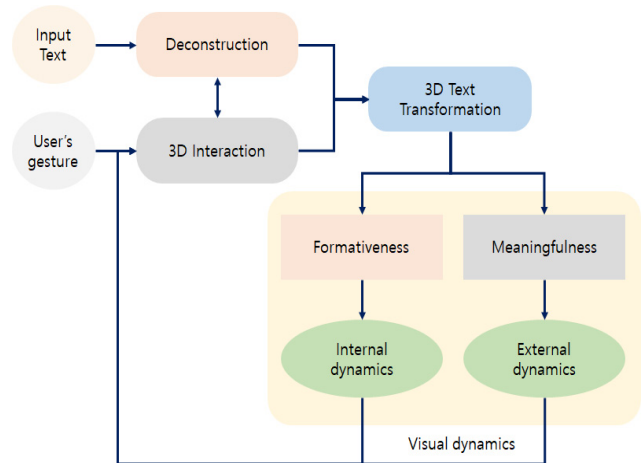


Fig. 5. The procedure of the proposed 3D text visualization tool

Sometimes text is a means of conveying meaning and sometimes it is an image. Text can generate spatial visibility by modifying its shape such as layout. The new visible space is related to the production of new meaning. In addition to the main purpose of information transmission, the control of the text form using the gesture of the viewer can make the viewer immerse in the artwork because it can form differentiated meaning for each individual. There are two transformations of text. One is a transformation of formativeness based on the visibility and spatiality of the text. Transformations in the text formality combined with the viewer's interaction create a new space with the creation of new external visual information. The other is the production of the internal meaning transformation. The result of the viewer's interaction is the link between the current textual representation and the viewer's past experiences. This association of time creates new meanings through the association of viewers and creates intangible internal space.

4.3 Exhibition

The developed 3D environmental text visualization system was applied for media art exhibition. In actual exhibition environment, sound information was added to increase the immersion feeling of viewer. The viewer has the copyright on the two most important elements of the exhibition environment- text and sound. In the exhibition, the sound is a variable element that changes according to the viewer's emotional state. 3D text, which generates bits according to music, can stimulate both the viewer's body and intent.

The resulting 3D text, which the viewer interacts with, moves along the x and y axes according to the gesture of the viewer and moves along the z axis according to the frequency of the sound. Fig. 6 is an installation diagram of 3D environmental text visualization system for exhibition.

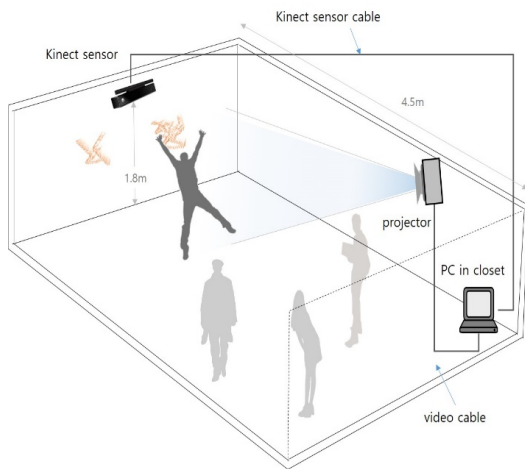


Fig. 6. Installation diagram of 3D Environmental Text Visualization system

The environment for the exhibition was implemented on a computer with Intel i5 CPU, 8G RAM, 256GB SSD, GeForce GTX graphics card, Windows 10 64 bit operating system. The text visualization application was developed using Visual Studio 2015 (including MFC), Microsoft Kinect for Windows SDK 2.0, OpenCV 3.2, and Openframeworks 0.9.8.

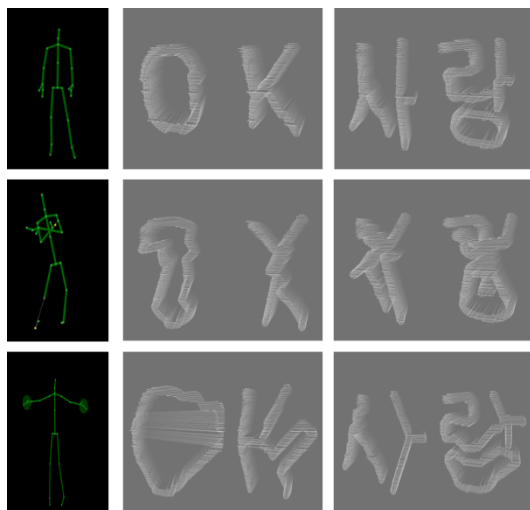


Fig. 7. Visual dynamics of text to the viewer's gestures

The developed text visualization tool provides two types of user interface for inputting text. Using the keyboard input, the viewer can select a desired language and fonts. A viewer who wants a more creative form of text input can use the touchpad. Our system has allowed the viewer to have a unique text input and familiarity through various input formats. Therefore, the viewer can perform the role of the creator of the content strongly requesting the interaction beyond the role of the content consumer.

Fig. 7 shows the various forms of visual dynamics produced by the gestures of the viewer. The viewer takes the gesture to remind or even express the different meanings of the same text if the form is different.

5. CONCLUSION

In this study, we proposed a visual dynamics model for expressing various human intention using text, and developed a method for visualizing the dynamics of text meanings based on the proposed model. The proposed 3D environmental text visualization tool is an interactive text authoring tool that provides an easy user interface and immediate and dynamic interaction. By allowing viewers to enter various types of input using a keyboard or a touch pad, it provides a flexible and free viewing environment. It is also a language-independent text visualization tool, which saves time and money in creating separate languages. Interaction with text is performed naturally by gesture recognition techniques. For the dynamic movement of the text synchronized with the gesture of the viewer, the deconstruction process of text was carried out. The viewer can feel the text alive through the deconstruction, and enjoy the dynamic space. Therefore, the proposed tool provides a particularly effective visualization in Asian languages that can be deconstructed into meaningful elements rather than English.

As a result of using the developed system in the exhibition, the viewer showed a great immersion and satisfaction in moving text like his avatar. The abundant distorted expressions of the deconstructed text maximize visual dynamics, creating an unexpected shape. Visual dynamics serves as a catalyst to reveal the meanings that have been latent in the viewer's unconsciousness, leading to various interpretations of the artwork from the viewer. The viewer creates an artistic 3D space by interacting with text to maximize his/her intention and emotions, produce various meanings, and show colorful and dynamic visual effects. Therefore, applying the developed system to the public art field of the future will give meaning to human life and actively change the environment.

REFERENCES

- [1] For the Guggenheim, <https://www.youtube.com/watch?v=OT62rD7wPkU/> (accessed Sep. 12, 2018).
- [2] Text Rain, <http://camilleutterback.com/projects/text-rain/> (accessed Sep. 12, 2018).
- [3] Written Forms, <http://camilleutterback.com/projects/written-forms/> (accessed Sep. 12, 2018).

- [4] R. Yoshioka, "3D Kanji: A new paradigm of 3D objects to exploit additional dimensions of human sense for enhancing expression," *Journal of Visual Languages & Computing*, vol. 28, pp. 250-272, 2015.
- [5] J. A. Russell, "Evidence of convergent validity on the dimensions of affect," *Journal of Personality and Social Psychology*, vol. 36, no. 10, pp. 1152-1168, 1978.
- [6] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *Journal of research in Personality*, vol. 11, no. 3, pp. 273-294, 1977.
- [7] H. Lövhheim, "A new three-dimensional model for emotions and monoamine neurotransmitters," *Medical hypotheses*, vol. 78, no. 2, pp. 341-348, 2012.
- [8] S. Y. Lim, "Emotional Communication on Interactive Typography System," *International Journal of Contents*, vol. 14, no. 2, pp. 41-44, 2018.
- [9] S. Y. Lim, *3D Interactive Typography with Character Deconstruction on Visual Dynamics Model*, Ph.D. Thesis, Kyungpook National University, Daegu, Republic of Korea, 2018.

**Sooyeon Lim**

She received Ph.D. in Computer Engineering and Digital Media Art from Kyungpook Nat'l Univ., Korea in 2006 and 2017 respectively. She is a professor who is working for Department of Game at Dongyang Univ. in Korea. Her research interests include Interactive Art,

HCI, 3D character animation and Image processing for diverse applications.