Content Enhanced Word Art with Depth Perception

Shao-Wei Yen[1] waybering@gmail.com
Chih-Kuo Yeh[2] simpson.ycg@gmail.com
Charles C. Morace[2] charles.c.morace@gmail.com
Sheng-Yuan Chen[2] chensyt@gmail.com
Shih-Syun Lin[3] catchylss@gmail.com
Chia-Hsiang Chen[2] traditionalttx@gmail.com
Tong-Yee Lee[2] tonylee@mail.ncku.edu.tw

[1]National Changhua University of Education, [2]National Cheng-Kung University, [3]National Taiwan Ocean University

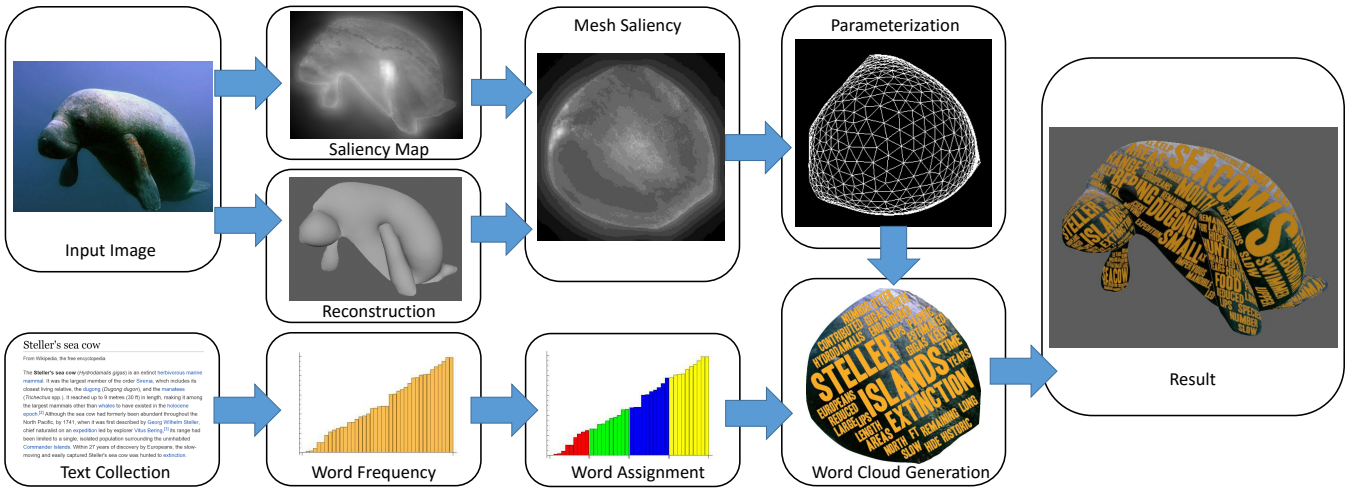

Figure 1: *An overview of the proposed system workflow and example result. A visual saliency map and high-relief 3D reconstruction [Yeh et. al. 2016] is obtained from an input image to direct the placement of words extracted from a collection of text.*

Keywords: word art, text visualization, eye tracking, saliency image, mesh parameterization

Concepts: •**Applied computing** → **Fine arts;**

1 Introduction

Word art has characteristics of both text and an image by combining words into a collage that depicts an associated shape or object. This type of artwork has the unique perceptual property that when seen as a whole, it is perceived similar to an image, but when viewing a part, it can be read word by word. Thus word art can serve a dual purpose, efficiently communicating a theme using shape, and accurately communicating meaning using text.

Manually creating word art is a laborious process. Therefore we propose a computational approach to assist artist and designers in the creation process. This work is closely related to the topic of word cloud generation. Previous research focuses on two-dimensional word clouds, where words of different size and color fill the interior of a simple shape based on word frequency and word significance, respectively. In contrast, this work introduces a new approach; we use visual saliency as a guide for the position of text in the word art, and texture-map words to a high-relief 3D reconstruction of an input image to preserve important depth cues in the result. The primary goals of our approach are to position words of greater significance in regions of greater visual saliency and to produce word art with perceivable depth.

SA '16, December 05-08, 2016, , Macao
ISBN: 978-1-4503-4540-8/16/12
DOI: http://dx.doi.org/10.1145/3005274.3005280

2 Technical Approach

The input to the proposed system is a collection of text and a related image. Initial preprocessing includes three independent steps. Firstly, we remove stop words and calculate word frequency from the collection of text. Secondly, we compute a visual saliency map which estimates the distribution of a viewer's focal attention over the input image. Thirdly, we use a 3D reconstruction method [Yeh et al. 2016] to generate high-relief triangle meshes for multiple regions of the input image.

To give the word art the appearance of depth, we generate word clouds as textures and map them to the surface of corresponding meshes. Each mesh in the reconstruction is parameterized using a conformal surface parameterization. This parameterization yields a bijective mapping between the 3D model space and 2D UV texture space which preserves local angles. To position words of greater significance in regions of greater visual saliency, we assign a saliency value to each mesh according to the saliency map computed from the input image. Some meshes in the reconstruction may be too small to display distinguishable words. Therefore meshes are assigned the original image as a texture if they a smaller than a user-defined threshold value. For each mesh larger than the threshold value size, we assign a subcollection of the text based on the mesh's visual salience value and use the word frequency statistics so that meshes with the greatest visual saliency are assigned the

most frequently used words.

With the surface parameterization and associated subcollection of words, a word cloud is generated in the UV space of each mesh by applying the method presented by [Chi et al. 2015]. This method is user-assisted, with the range of word size and word orientation being controlled by the user. Finally, the generated word cloud is texture mapped onto the surface of the corresponding 3D mesh using the previously determined parameterization.

3 Limitation and Future Work

One drawback to the proposed method is that the saliency map and the 3D reconstruction may have little correspondence with each other. In the future, we plan to integrate the saliency model with the 3D reconstruction so that connected regions in the saliency model remain connected in the reconstruction. We also plan to verify the quality of our results with an extensive user study.

Acknowledgements

This work was supported in part by the Ministry of Science and Technology, Taiwan (contracts MOST-105-2811-E-006-016, MOST-104-2221-E-006-044-MY3 and MOST-103-2221-E-006-106-MY3).

References

CHI, M.-T., LIN, S.-S., CHEN, S.-Y., LIN, C.-H., AND LEE, T.-Y. 2015. Morphable word clouds for time-varying text data visualization. *IEEE Transactions on Visualization and Computer Graphics 21*, 12 (Dec), 1415–1426.

YEH, C.-K., HUANG, S.-Y., JAYARAMAN, P., FU, C.-W., AND LEE, T.-Y. 2016. Interactive high-relief reconstruction for organic and double-sided objects from a photo. *IEEE Transactions on Visualization and Computer Graphics 99*, 1–1.