

Automatic Advertisement Copy Generation System from Images

Koichi Yamagata^{*1}, Masato Konno^{*1}, Maki Sakamoto^{*1}

^{*1} Graduate School of Informatics and Engineering, The University of Electro-Communications

When we want to sell something, the presence of a good advertisement copy often affects sales. In this research, we develop an automatic advertisement copy generation system. Most existing systems only enter keywords, and the potential image of the product is not necessarily reflected by keywords only. We propose a method to generate advertisement copies using images as input to convey potential messages and the world view. This method uses Word2vec and color information of ad images, and both were confirmed to be effective by evaluation experiments. In the evaluation experiments, the mean score of the proposed method was significantly larger than 4 out of 7, and most of the subjects answered positively to ad copies of our method.

1. Introduction

Advertisement copies describe the features of products with a short number of characters and impact sentences, which are factors that greatly contribute to building brands such as companies and promoting purchase willingness of products. In recent years, there have been many researches on the generation of Japanese sentences and advertisement copy generation. In existing researches, however, they did not focus on describing potential messages and the view of the world that the producer wanted to convey.

In this research, we propose a method to generate appropriate copies reflecting color and sensitivity of given images by using a database that maintains the relationship between color and words.

There are also several prior studies on the relationship between color and sensitivity. Iiba et al. [Iiba 13] focused on the relationship between color and word sensibility, and constructed a system that recommends appropriate colors and fonts for textual sensibility images considering the effect reminiscent of the color of words. Further, Nakamura et al. [Nakamura 12] focused on the relationship between color and lyrics and constructed a music retrieval system with color as input. In this research, we focus on the relationship between colors and words and aim to automatically generate advertisement copies with color input.

Recently, in the field of natural language processing, various methods handling words as distributed expressions are increasing. Mikolov et al. [Mikolov 13] developed Word2vec which is a new model that improves the precision of vector operation of words considering the similarity between words in sentences. In this research, we utilize Word2vec to extract synonyms from a large corpus.

Existing researches on advertising copying were mainly methods of inputting words and keywords, and they were not taken into consideration to convey the latent message or the world view that the creator wished to convey. On the other hand, in this research, we construct an automatic advertisement copy generation system using images as inputs. Color information is extracted from the input image, and a large number of advertisement copies are generated by using a database that

maintains the relationship between words and colors, from which adequate copies are determined as outputs.

2. Methods

The flow of the proposed method is as follows:

1. Color information contained in given image is extracted, and some keywords are extracted from the database storing the relationship between words and colors.
2. From the lyrics database, some phrases including the keywords extracted in step 1 are searched and extracted as sentence templates.
3. By using a deep neural network (dnn) based classification model that classifies words used for ad copies or not trained by ad copies corpus, word candidates are selected from the word extracted in step 1.
4. The noun in the template sentences extracted in step 2 are replaced by the words selected in step 3, and they are taken as copy candidate sentences.
5. The ad copy candidate sentences are evaluated by the similarities given by Word2vec, and the most evaluated ad copy candidate sentence is outputted.

In step 1, we utilize the word-color database constructed by [Konno 18]. This database was constructed based on the idea of the relation between sensitivities and music/colors studied by [Nakamura 11]. In this database, among the songs released between 1968 and 2017, we defined words that were often used for lyrics as primitive words (PWs), and each PW has a 45-dimensional color information vector obtained by a psychological experiment. Figure 1 shows 45 colors to define the 45-dimensional color vector space. For words other than PWs that were not psychologically tested, latent semantic analysis was performed on them, and color information vectors were given based on the similarities with PWs.

In step 3, pytorch is used as a dnn library. The input data is a bag-of-words vector of a sentence to be classified, and the output data is “copy” or “not-copy”.

In step 4, we use Word2vec model which learned 100,000 songs of Japanese lyrics corpus, and which contains similarities with a poetic point of view. Using this model, we calculate the similarities between words (nouns, verbs, adjectives etc.) contained in sentences and evaluate candidate sentences. The results of evaluation are used to select the most appropriate ad copies.

Contact: Koichi Yamagata, UEC,
1-5-1, Chofugaoka, Chofu, Tokyo 182-8585, Japan,
koichi.yamagata@uec.ac.jp, 042-443-5535

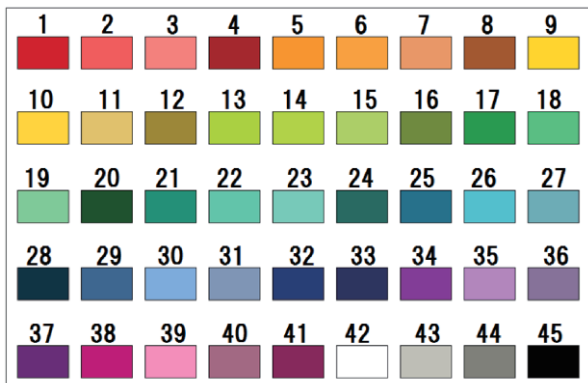


Fig. 1: 45 colors to define the 45-dimmmensional color vector space.

3. System Evaluation

In order to evaluate our system, we conducted an evaluation experiment with 30 subjects (7 females and 23 men, mean age = 22.63). We adopted 20 advertisement images actually used from four categories (beverage and food, travel, beauty, fashion), and we prepared five ad copies for each advertisement image by five methods:

- A) original ad copy
- B) proposed method
- C) proposed method without Word2vec
- D) proposed method without using the image
- E) RNN

For example, when input image was a picture of three can coffees against the background of the night sky, an ad copy of method B (translated into English) was “You will surely drink it tonight”, method C outputted “You laugh, even the deadly midnight dropping the mountain, forever”, method D outputted “Eat coffee and coffee”, method E outputted “Feel coffee and go”.

Subjects were asked to evaluate a total of 100 advertisement copies for the images including the original ad copies and the ad copies outputted by the proposed method. The subjects evaluated the following questionnaires on a scale from 1 to 7:

- i. Is it appropriate as an ad copy in this category?
- ii. Is it appropriate as an ad copy regardless of category?
- iii. Is the grammar of the ad copy appropriate?
- iv. Does the ad copy follow the impression of the image?

Table 1 shows the results of the questionnaires (averages and standard errors of scores). We can see that the mean score of the proposed method B is much larger than 3, and most of the subjects answered positively to our method. Comparing the methods A and B, we can see that the proposed method is slightly less than the scores of original advertisements. However, it is not so bad and the difference is small. In the food category, it is confirmed that the proposed method has smaller score differences with respect to the original advertisement. Comparing the methods B and C, we

can see that the use of Word2vec is effective. Comparing the methods B and D, we can see that the use of images is effective. Comparing the methods B and E, we can see that the proposed method is much better than the RNN method.

Table 1: Results of questionnaires to evaluate each method (averages and standard errors of scores)

	i	ii	iii	iv
A	5.29 ± 0.07	5.75 ± 0.06	6.03 ± 0.06	5.37 ± 0.07
B	4.90 ± 0.07	5.17 ± 0.06	5.55 ± 0.06	4.22 ± 0.08
C	2.82 ± 0.07	3.15 ± 0.08	3.31 ± 0.08	2.51 ± 0.06
D	4.23 ± 0.08	3.92 ± 0.08	3.81 ± 0.09	3.20 ± 0.07
E	3.18 ± 0.07	2.74 ± 0.07	2.05 ± 0.06	2.75 ± 0.07

4. Conclusion

This study proposed a method to generate advertisement copies using images as input to convey potential messages and the world view. This method uses not only Word2vec but also color information of images, and both were confirmed to be effective by evaluation experiments. In the evaluation experiments, the mean score of the proposed method was significantly larger than 4 out of 7, and most of the subjects answered positively to our method. The proposed method was slightly less than the scores of original advertisements. However, it was not so bad and the difference was small.

References

- [Iiba 13] Iiba, S., Doizaki, R., and Sakamoto, M.: Color and Font Recommendations based on Mental Images of Text, Transactions of the Virtual Reality Society of Japan, 18(3), pp.217-226 (2013)
- [Nakamura 12] Nakamura, T., Utsumi, A., and Sakamoto, M.: Music Retrieval Based on the Relation between Color Association and Lyrics, Transactions of the Japanese Society for Artificial Intelligence, Volume 27, Issue 3, pp. 163-175, 2012.
- [Mikolov 13] Mikolov, T., Yih, W., Zweig, G.: Linguistic Regularities in Continuous Space Word Representations, Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT-2013) (2013)
- [Konno 18] Konno, M., Suzuki, K., and Sakamoto, M.: Sentence Generation System Using Affective Image, 2018 Joint 10th International Conference SCIS and 19th ISIS, 678-682.
- [Nakamura 11] Nakamura, T., Kawanishi, K., and Sakamoto, M.: A Possibility of Music Recommendation Based on Lyrics and Color, Transactions of the Japanese Society for Artificial Intelligence, Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Volume J94-A, pp.85-94, No.2 (2011)