Engineering Applications of Artificial Intelligence 73 (2018) 149–162

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Extracting and summarizing affective features and responses from online product descriptions and reviews: A Kansei text mining approach

W.M. Wang [a], Z. Li [a,*], Z.G. Tian [a], J.W. Wang [a], M.N. Cheng [b]

[a] Guangdong Provincial Key Lab. of Computer Integrated Manufacturing Systems, School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou, Guangdong, 510006, China

[b] Knowledge Management and Innovation Research Centre, Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, 999077, China

ARTICLE INFO

Keywords:
Affective mining
Opinion mining
Customer reviews
Affective design
Kansei engineering

ABSTRACT

Today's product design takes into account the affective aspects of products, such as aesthetics and comfort, as much as reliability and physical quality. Manufacturers need to understand the consumers' affective preferences and responses to product features in order to improve their products. Conventional approaches use manual methods, such as questionnaires and surveys, to discover product features and affective preferences, and then correlate their relationships. This is one-time, labour-intensive, and time-consuming process. There is a need to develop an automated and unsupervised method to efficiently identify the affective information. In particular, text mining is an automatic approach to extract useful information from text, while Kansei engineering studies product affective attributes. In this paper, we propose a Kansei text mining approach which incorporates text mining and Kansei engineering approaches to automatically extract and summarize product features and their corresponding affective responses based on online product descriptions and consumer reviews. Users can efficiently and timely review the affective aspects of the products. In order to evaluate the effectiveness of the proposed approach, experiments have been conducted on the basis of public data from Amazon.com. The results showed that the proposed approach can effectively identify the affective information in terms of feature–affective opinions. In addition, we have developed a prototype system that visualizes product features, affective attributes, affective keywords, and their relationships. The proposed approach not only helps consumers making purchase decisions, but also helps manufacturers understanding their products and competitors' products, which might provide insights into their product development.

1. Introduction

Today customers no longer buy products, they buy experiences (Norman, 2004). Therefore, today's product design not only takes into account the reliability and physical quality, but also takes into account the affective aspects of the product to meet the emotional needs of consumers and improve consumer satisfaction (Rosler et al., 2009). The affective aspects of the products are studied by researchers and manufacturers based on consumer-centred design (Nagamachi and Lokman, 2010). The researchers investigate the qualitative demands of consumers by assessing their psychological feedbacks after they used the products. The manufacturers collect and analyse these feedbacks and apply the results of the analysis to their production plans (Vieira et al., 2017). In particular, Kansei engineering (or affective engineering) is a mechanism for translating human emotional needs into product design elements quantitatively (Nagamachi, 1989; Nagamachi and Lokman, 2016). Kansei is a Japanese term that represents emotions and impressions. In Kansei engineering studies, surveys are always used to study the relationship between affective attributes and product design features (Llinares and Page, 2011). The most commonly used method is the semantic differential (SD) method, which is a rating scale used to measure respondents' opinions and attitudes towards a given object (Osgood et al., 1957). Researchers use the SD method to design questionnaires to measure subjective consumer impressions of the product (Yan et al., 2008). The questionnaire consists of a list of words called Kansei attributes. Each Kansei attribute refers to a particular emotional expression (Chou, 2016). Each Kansei attribute consists of a bipolar pair of Kansei words (i.e. a positive word and a negative word, such as beautiful–ugly) (Friborg et al., 2006). Respondents are usually asked to rate an N-point scale between the bipolar words to represent the subjective assessment of the product.

* Corresponding author.
E-mail address: piersli@foxmail.com (Z. Li).

https://doi.org/10.1016/j.engappai.2018.05.005
Received 25 October 2017; Received in revised form 5 February 2018; Accepted 15 May 2018

The conventional survey-based approach provides high-quality affective data, which has been widely used in many affective design studies (e.g. Yan and Nakamori, 2010; Chou, 2016; Kwong et al., 2016). However, they rely on users to actively participate in the study, therefore, most of the existing studies are carried out in a relatively small scale of operation. For example, Chou (2016) involved seven users in their Kansei evaluation of 10 products; Jiang et al. (2015a) involved four users to evaluate 10 products; Guo et al. (2016) studied 36 people in 16 designs. Moreover, the survey questions of traditional method are designed based on expert thinking rather than customers' point of view (Hsiao et al., 2017). Respondents are only able to passively respond to the expert designed questions. In addition, respondents may not be the consumers of the target products. Furthermore, due to the time-consuming and labour-intensive process of questionnaire design, distribution and collection, the survey-based approach is inadequate to involve too many users and products. It is also not suitable to be conducted in real-time basis. However, due to the high industry competition and the product customization trend, more and more products are launched to the market in a very short period of time. Therefore, it is necessary to develop an automated method to efficiently review the consumers' affective feedbacks.

Text mining refers to the use of techniques in natural language processing, computational linguistics, and statistical analysis to systematically and automatically identify and extract useful information from texts (Liu, 2012). Recently, the mining of useful information from online product data has received much attention in many areas (Jin et al., 2015). Most online shopping sites allow consumers to provide their product reviews after purchasing a product. It provides direct, real-time, and verified data from the consumers' perspective. Due to the massive amount of online consumer reviews, the application of text mining is promising to extract important affective information from the consumers' perspective in an efficient and effective manner. In particular, sentiment analysis is a subarea of text mining. The existing studies use sentiment analysis to extract emotional information (Liu, 2012). However, most of them only divide the information into three states: positive, negative and neutral (e.g. Vilares et al., 2017), which is not enough for affective product design.

In this paper, we aim to combine Kansei engineering and text mining approaches to develop Kansei text mining approach which uses text mining to automatically convert unstructured product-related texts to feature–affective opinions. The main contributions of this paper are as follows: (1) We propose an automatic and unsupervised text mining method that combines the information of online product descriptions with consumer reviews to extract product features as well as their corresponding consumer affective responses. (2) We classify affective opinions into multiple affective attributes and relate to product features. Compared with the existing sentiment analysis methods, they mainly divide an opinion into positive, negative and neutral from a single perspective (i.e. either good or bad). (3) We propose a semi-automatic method to select generic Kansei words and attributes based on publicly available data. Therefore, the results could be reused and applied to other products. (4) We propose a summarization method to summarize the relationship between product features and consumer affective responses. We also design and develop a prototype system to visualize the summaries.

We organize the rest of this paper as follows. Section 2 presents a review of the related studies. Section 3 describes the proposed approach. An experiment has been conducted to evaluate the proposed approach. The experimental results are discussed in Section 4. Section 4 also describes the application of the proposed approach and the development of the prototype system. Lastly, Section 5 provides conclusions, limitations, and recommendations for further work.

2. Related studies

Kansei engineering is a product development method used to investigate human feelings and to discover quantitative relationships between the affective responses and design features (Nagamachi, 1989; Nagamachi and Lokman, 2016). By using Kansei engineering, a lot of research has been done to improve product and service design. However, the data collection method of the existing studies is very similar. Most of them manually collect customized Kansei words from various data sources, such as customer interviews, expert interviews, journal articles, magazines, news, advertisements, and more. They then use the collected Kansei words and the SD method to design the Kansei questionnaire. Finally, the questionnaire is distributed to a group of target respondents to collect their emotional feedback. For instances, Chan et al. (2011), Fung et al. (2014) and Jiang et al. (2015a, b) used this method to conduct a set of customer surveys on the affective design studies of mobile phones. Llinares and Page (2011) designed a questionnaire that used this method to measure subjective consumer perceptions that influence property purchase decisions. Shieh et al. (2016) used this method to explore the relationship between the shape and colour of toothbrushes. Li and Yan (2016) used this method to analyse the relationships among service attributes, Kansei words, and customer satisfaction of hotel services.

On the other hand, due to the advancement of information technology, more and more customer data is available on the Internet. Researchers can use text mining to analyse online consumer reviews to collect useful information (Liu, 2012). Text mining is a process involving the use of natural language processing and machine learning to obtain high-quality information from unstructured text (Feldman and Sanger, 2007). A typical text mining process begins with pre-processing. In general, it uses natural language processing techniques to perform pre-processing. It includes sentence segmentation, tokenization, and part-of-speech (POS) tagging, etc. Sentence segmentation is a process to divide a text into paragraphs and sentences. Tokenization is a process of converting a text into tokens (i.e. words). POS tagging is a process of assigning a POS to a word. Text mining then uses different analytical methods, such as rules, statistical methods or data mining methods, to discover interesting patterns. Finally, post-processing may be applied to interpret and represent the analysed results in different formats, such as graphics or mappings. In particular, sentiment analysis (sometimes also known as opinion mining) focuses on identifying, extracting, and quantifying the writer's affective attitudes towards the subject based on his/her written text (Yadollahi et al., 2017). Sentiment analysis has been applied to a variety of applications, including learning and education (Ortigosa et al., 2014), healthcare (Desmet and Hoste, 2013), finance (Oliveira et al., 2016), customer relationship management (Kang and Park, 2014), and so on. A number of studies have been conducted to analyse the sentiment of online reviews. For instances, Liu et al. (2017) proposed a method based on the sentiment analysis and fuzzy set theory to rank products by online reviews. The method determined the positive, neutral or negative sentiment orientation of each review, and then constructed an intuitionistic fuzzy number to represent the performance of the product regarding its product feature. In the study of Zhou et al. (2017), a combination of affective lexicons and a rough-set technique is proposed to predict sentence sentiments of individual product features, thereby augmenting a feature model by integrating positive and negative opinions of consumers. Cho et al. (2014) proposed a data-driven approach that adapts multiple sentiment dictionaries to different domains. They use the ratio of the positive/negative training data and remove entries that have no contribution to the classification.

Contrary to existing research, we propose a semi-automatic method to collect a generic set of Kansei words and attributes from publicly available data. Second, we propose an automatic and unsupervised text mining method to extract product features from online product