Mining of affective responses and affective intentions of products from unstructured text

W. M. Wang[a], Z. Li[a,*], Layne Liu[a], Z. G. Tian[a], Eric Tsui[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

[*] Corresponding author

Abstract

The current product design not only takes into account the function and reliability, but also concerns about the affective aspects to meet the emotional needs of consumers. However, there is always a gap between the affective intention of the manufacturer and the affective response of the consumer. In order to improve consumer satisfaction, it is necessary to understand this gap. Traditional methods rely on manual surveys to understand the gap, which are costly, time consuming and small in scale. There is a need to develop an automated method to efficiently extract the manufacturer's intentions and the consumer's responses in terms of affective aspects. In the past few years, big data is gaining more and more interest. People analyze big data to obtain useful information for marketing and product design. Manufacturers provide online products descriptions, and consumers provide online reviews after purchasing a product. Handling these data is expected to understand affective information of the product from the perspectives of manufacturers and consumers. In this paper, we propose a text mining method to extract the affective intentions and the affective responses and of products from the product description and consumer reviews, respectively. We build an affective profile for each product. And based on the affective profile, we can visualize the gap between the consumer's affective responses and the manufacturer's affective intentions of the product. We use Amazon.com data to conduct a case study to study the effectiveness of the proposed method. We construct the affective profiles for selected products and analyze the gap between the manufacture's affective intention and consumer's affective responses to the products. We also introduce the use of affective information in product recommendations.

Keywords

Text mining; Affective response; Affective intention; Affective profile; Product recommendation; Affective design

1. Introduction

Products with good affective design can enhance consumer satisfaction. Today, more and more manufacturers take into account the affective aspects of the product, such as aesthetics and comfort, as much as properties like reliability and physical quality (Rosler et al. 2009). The manufacturers investigate the psychological feelings and needs of the consumer in order to apply them in to the production plan (Vieira et al., 2017). In particular, Kansei is a Japanese, meaning sensibility, impression, and emotion. And Kansei engineering or affective engineering is a mechanism for translating consumer emotional needs into product design elements quantitatively (Nagamachi, 1989; Nagamachi & Lokman, 2016). At the early stages of the Kansei project, surveys are always used to study the relationship between affective attributes and design elements (Llinares & Page, 2011). The traditional method is based on the semantic differential method (Osgood et al., 1957) to design the questionnaire. The questionnaire consists of a list of Kansei attributes. Each Kansei attribute refers to the representation of emotional connotation (Chou, 2016). They are product attributes commonly used to obtain subjective impressions of the manufacturers and consumers on the product (Yan et al., 2008). 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). The questionnaires are assigned to a group of consumers. The consumers use an N-point scale to represent a subjective assessment of a particular product or service. This approach has been widely used in many affective design studies (e.g. Yan & Nakamori, 2010; Chou, 2016).

Conventional methods provide high-quality affective data, but only in a relatively small scale of operation. For example, Chou (2016) involved 7 users in their Kansei evaluation of 10 products; Jiang et al. (2015a) involved 4 users to evaluate 10 products; 36 participants participated in 16 designs in the study of Guo et al. (2016). A larger scale of affective design survey has been done by Hsiao et al. (2017). They used a web-based survey, which collected 118 valid questionnaire for a single service. In addition, most of the existing research focuses on investigating the relationships between affective attributes and design attributes. However, there is always a gap between the affective intentions of the manufacturers and the affective responses of the consumers (Hsu et al., 2000). In order to effectively improve consumer satisfaction, it is necessary to understand this gap in order to reduce this gap.

Recently, big data in the information technology, marketing and manufacturing, and many other areas have received a lot of attention. There is currently a lot of information online. The manufacturer provides an online product description of its products, reflecting its perception of the products. After the purchase of products, consumers offer their views through product reviews. The processing of these data is expected to understand the affective aspects of the product from the manufacturer's perspective and the consumer's perspective.

In this paper, we aim to use text mining to automatically convert unstructured product-related texts into affective-related information. The main contributions of this paper are as follows: 1) we propose an automatic text mining method extract the affective information of the product from unstructured text; 2) we assign classify the affective information into different Kansei attributes and assign a degree to each attribute to construct the affective profile of the product and the consumer; 3) we examine the gap between the consumer's affective responses and the manufacturer's affective intention through a case study; and 4) we evaluate the performance of product recommendation using affective information. We evaluate the effectiveness of using affective information in product recommendation.

We organize the rest of this paper as follows. First, Section 2 presents a review of the related studies. Section 3 describes the proposed method. Section 4 describes the application of the proposed method in the automotive product gap analysis and product recommendation. The results and discussions are also discussed in Section 4. Finally, Section 5 provides conclusions and recommendations for further work.

2. Related studies

A lot of research has been done using the Kansei engineering to improve product and service design. For example, Chan et al. (2011) developed a fuzzy regression method to identify the non-linear and fuzzy relationships between affective responses and design variables. Llinares et al. (2011) proposed the use of Kano's model to analyze the impact of different subjective attributes on consumer purchasing decisions based on semantic differential and regression analysis. Li and Han (2016) used the Kansei Engineering to study the relationship between the service attributes, Kansei, and customer satisfaction of hotel services. Shieh et al. (2016) explored the relationship between the shape and color of the toothbrush by combining the Kansei engineering and rough set theory. Fung et al. (2014) proposed a guided search genetic algorithm approach based on mining rules to determine the optimal design attribute setting for affective design. Guo et al. (2016) proposed a method based on semantic differential method, back propagation neural network and genetic algorithm to extract user-centered emotional dimensions, identify the quantitative relationship between key design factors and emotional dimension, and find a near-optimal design. An adaptive neuro-fuzzy inference system based on rough set and particle swarm optimization is proposed by Jiang et al. (2015b) to model customer satisfaction with affective design and further improve the modeling accuracy.

Kwong et al. (2016) proposed a fuzzy regression approach for modeling customer satisfaction and developing cost models, a chaos-based fuzzy regression approach for generating product utility functions, and a non-dominated sorting genetic algorithm-II for solving multi-objective optimization.

In particular, Hsu et al. (2000) have a similar objective with this paper. They used semantic differential to study the differences of perception between designers and users. They used 24 real telephone samples, 20 designers and 20 users. While in Hsiao et al. (2017), they used text mining to identify the service elements and Kansei words. Contrast to these studies, we use consumer reviews and product descriptions as the source of subjective assessment. Second, we automatically extract affective attributes and measure the degrees of the attributes in affective analysis. Third, we analyze the intention of the manufacturers who design and sell the product rather than analyzing the views of general designers. Lastly, we use fuzzy set theory to measure linguistic uncertainty.

3. Affective profile construction

The architecture of the proposed affective profile construction method is shown in Figure 1. It consists of 4 main steps including Kansei attributes and words collection, single-document affective analysis, product affective analysis, and applications.

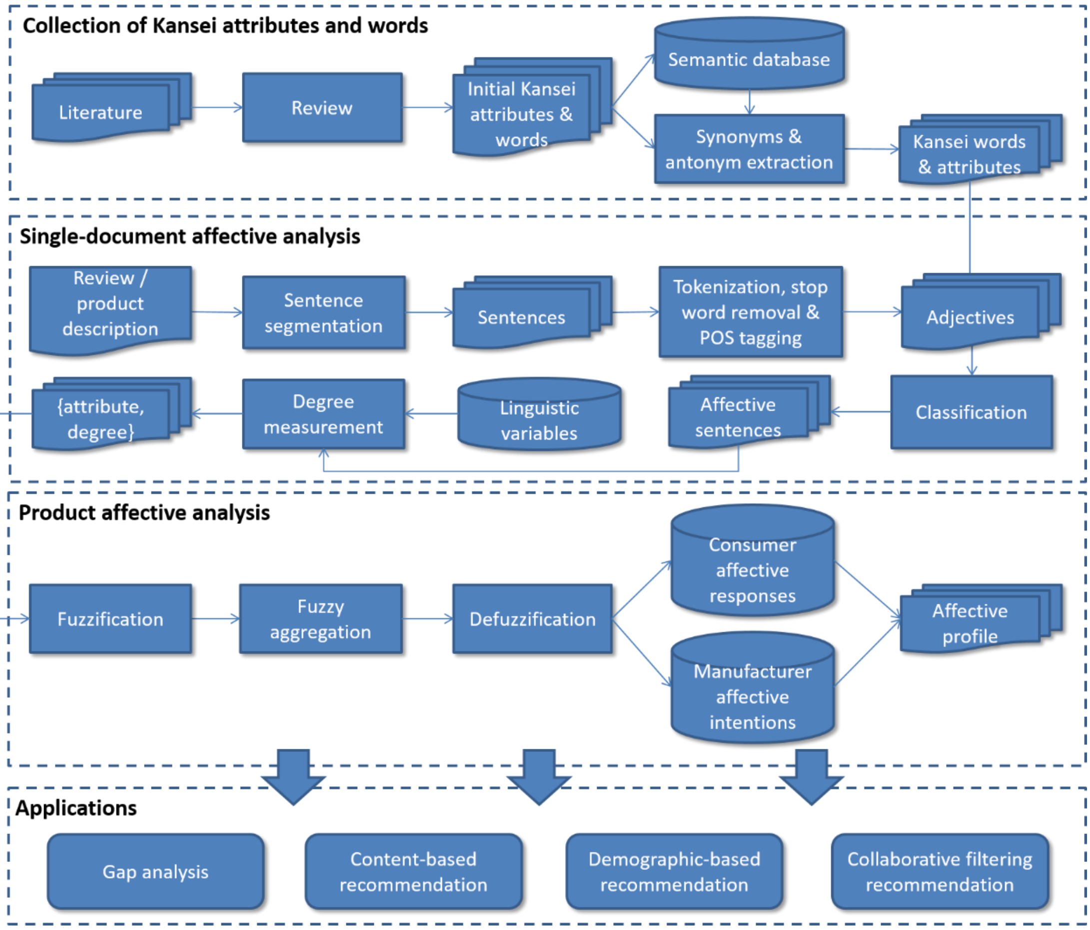


Figure 1 Architecture of affective profile construction

3.1 Collection of Kansei attributes and words

Kansei attributes and Kansei words are usually identified by brainstorming and interviews (e.g. Vieira et al., 2017; Grimsæth, 2005), manually extracted through reviews (e.g. Chou, 2016; Shieh et al., 2016), and Kansei clustering (Huang et al., 2012). In particular, Kansei clustering classifies Kansei words into clusters of Kansei attributes based on fuzzy logic and manual evaluation (Huang et al., 2012).

In this study, we use a semi-automatic method to extract the Kansei attributes and Kansei words. We first manually select the Kansei attributes and Kansei words that are commonly used in the literature. Then, we use WordNet, which is a semantic database commonly used in text processing research, to find synonyms and antonyms of the Kansei words as the additional Kansei words. As described in the literature, most Kansei words are adjectives (e.g. Boran et al., 2014; Shieh et al., 2016). Therefore, we use WordNet's adjectives dictionary (Fellbaum, 1998) for the extraction.

Based on this method, we identify 15 Kansei attributes from the literature, including appealing, simple, comfortable, technological, soft, compact, reliable, unique, fashionable, precious, handy, pleasant, innovative, practical and lustrous. The Kansei words extracted from the literature and WordNet are shown in Table 1.

Table 1 Kansei attributes and words

Attributes		**Words from literature**	**Words from WordNet**	**Word count**	**References**
appealing	+ve	aesthetic, appealing, artistic, cute, elegant, exquisite, eye-catching, good-looking	e.g. beautiful, attractive, fine-looking	284	Chou, 2016; Guo et al., 2015; Jiao et al., 2006; Shieh et al., 2016; Barone et al., 2007; Llinares & Page, 2011
	-ve	inaesthetic, artless	e.g. unattractive, ugly, unappealing	211	
simple	+ve	simplificative, plain, simple	e.g. easy, effortless, uncomplicated	273	Chou, 2016; Fung et al., 2014; Guo et al., 2015; Jiao et al., 2006; Hsiao et al., 2017; Llinares & Page, 2011
	-ve	complex, complicated, dazzling, comprehensive	e.g. mazy, multiplex, composite	601	
comfortable	+ve	comfortable, handling comfort, cosy	e.g. comforted, relaxed, free	95	Jiao et al., 2006; Barone et al., 2007; Llinares & Page, 2011
	-ve	restrained	e.g. cautious, uncomfortable, unemotional	108	
technological	+ve	technological, hi-tech	e.g. high-tech, advanced, sophisticated	60	Chou, 2016; Fung et al., 2014; Llinares & Page, 2011
	-ve	classic, classical	e.g. low-tech, neoclassic, neoclassical	74	
soft	+ve	soft, smooth	e.g. lithe, gentle, feeble	454	Chou, 2016; Guo et al., 2015; Vieira et al., 2017; Bahn et al., 2009
	-ve	hard	e.g. stiff, firm, strong	646	
compact	+ve	compact, delicate, concise, refined	e.g. terse, fine, small	527	Chou, 2016; Guo et al., 2015; Bahn et al., 2009; Llinares & Page, 2011
	-ve	loose, coarse, sloppy	e.g. lax, unbound, reckless	590	
reliable	+ve	quality, high-quality,	e.g. correct,	533	Chou, 2016; Guo et al., 2015;

		reliable, sturdy, safe, accurate, robust, solid, durable	precise, exact		Jiao et al., 2006; Hsiao et al., 2017; Vieira et al., 2017; Bahn et al., 2009; Barone et al., 2007; Llinares & Page, 2011
	-ve	unreliable	e.g. inaccurate, imprecise, incorrect	574	
unique	+ve	unique, personalized, distinguished, particular, rare, tailor-made	e.g. individual, scarce, singular	389	Chou, 2016; Fung et al., 2014; Guo et al., 2015; Shieh et al., 2016; Llinares & Page, 2011
	-ve	general, common	e.g. popular, frequent, public	436	
fashionable	+ve	contemporary, stylish, fashionable, futuristic, modern, youthful	e.g. current, new, pop	174	Chou, 2016; Jiao et al., 2006; Hsiao et al., 2017; Barone et al., 2007; Llinares & Page, 2011
	-ve	traditional	e.g. conventional, old, past	146	
precious	+ve	precious, luxurious, luxury	e.g. expensive, valuable, big-ticket	115	Chou, 2016; Hsiao et al., 2017; Shieh et al., 2016; Llinares & Page, 2011
	-ve	low-cost, concessional	e.g. cheap, inexpensive, cut-price	91	
handy	+ve	handy, portable, ingenious	e.g. accessible, convenient, approachable	98	Chou, 2016; Fung et al., 2014; Jiao et al., 2006
	-ve	bulky	e.g. inaccessible, unaccessible, outback	133	
pleasant	+ve	enjoyable, cheerful, delightful, like, pleasant, peaceful	e.g. clam, quite, still	164	Jiao et al., 2006; Shieh et al., 2016; Vieira et al., 2017; Llinares & Page, 2011
	-ve	dislike, unpleasant, oppressive	e.g. furious, angry, unpeaceful	353	
innovative	+ve	novel, interesting, fresh, stimulating, innovative	e.g. revolutionary, originative, progressive	207	Chou, 2016; Guo et al., 2015; Jiao et al., 2006; Hsiao et al., 2017; Llinares & Page, 2011
	-ve	boring	e.g. dull, bored, unstimulating	277	
practical	+ve	practical, convenient, efficient	e.g. concrete, tangible, useful	120	Guo et al., 2015; Hsiao et al., 2017; Llinares & Page, 2011
	-ve	useless	e.g. conceptual, inefficient, inconvenient	115	
lustrous	+ve	lustrous, bright, light,	e.g. lighted, shiny,	131	Chou, 2016; Guo et al., 2015;

		glossy	shining		Hsiao et al., 2017; Llinares &
	-ve	dim	e.g. dark, dusky, lightless	128	Page, 2011

3.2 Single-document affective analysis

In the affective analysis, a set of documents related to the target products is collected. In the process of single-document affective analysis, an unstructured text is first divided into sentences by sentence segmentation based on detection of punctuations by regular expression. We then use natural language processing tool to perform tokenization, stop-word removal, and part-of-speech (POS) tagging. Tokenization is a process of converting a text into tokens (i.e. words). Stop-word removal is a process of deleting common words, such as pronoun, article, etc. In this paper, a stop word list is used to filter out frequently used words. POS tagging is a process of assigning a POS to a word. The adjectives of each sentence are extracted and mapped according to the Kansei words collected in the previous process. If it matches, the sentence is classified to have the corresponding Kansei attribute. Since a Kansei word may belong to multiple Kansei attributes and a sentence may consist of multiple Kansei words, a sentence may belong to multiple Kansei attributes.

After the Kansei attribute classification, the Kansei attribute is also measured based on a linguistic format scale with 9 degrees. It is adapted from the work of Chou (2016) which he used 7 degrees. The 9 degrees are “extremely”, “very”, “is”, “slightly”, “neutral”, “slightly not”, “is not”, “not very”, and “extremely not”. In order to automate the measurement of the degree of Kansei attributes, we extract the adverbs from the affective sentences. Then, the extracted adverbs are compared with a list of degree words. We use WordNet to identify the synonym for the degree words to construct the word list. Table 2 shows the degree words used in this study. In particular, if a sentence does not contain any degree word, it is classified as “is”. If a sentence contains the word “not” (or its synonym), the degree of Kansei attribute changes from a positive value to a negative value or from a negative value to a positive value. Based on the word list, each sentence containing the Kansei word is indexed with one or more than one degree. Thus, each sentence of the text can be converted into a set of Kansei attributes and their corresponding degrees. The following is an example to single-document affective analysis process:

Given a review that consists of 2 sentences s_1 and *s_2:*

This is a highly comfortable product to eliminate some stubborn odors from your Car Air conditioning system. But it is not elegant.

Matched Kansei words:

s_1: comfortable, s_2: elegant

Associated Kansei attribute:

s_1: comfortable, s_2: appealing

Extracted adverb and keyword:

s_1: highly, s_2: not

Mapped degree:

s_1: extremely, s_2: is not

Output:

s_1: {comfortable, extremely}, s_2: {appealing, is not}

Table 2 Words used for measuring the affective degree

Degree	**Words**
extremely (not)	extremely, boiling, bloody, damn, all-fired, all-firedly, fabulously, fantastically, incredibly, exceedingly, super, deathly, drop-dead, madly, insanely, deadly, deucedly, devilishly, inordinately, extraordinarily, brilliantly, highly, infernally, hellishly, positively, excellently, magnificently, splendidly, famously, wonderfully, wondrous, superbly, toppingly, marvellously, terrifically, marvelously, intensely, enormously, tremendously, hugely, staggeringly, big, goddam, goddamn, goddamned, piercingly, bitterly, bitingly, bitter, sharply, piping, steaming, precious, preciously, roaring, shockingly, stiff, whacking, whopping
(not) very	very, strongly, badly, bad, really, real, rattling, fucking, much, a lot, lots, a good deal, a great deal, fine, alright, all right, ok, hard, far, way, right smart, most, so, strongly, ever, dramatically, heaps, wholly, entirely, completely, totally, all, altogether, whole, right
slightly (not)	slightly, somewhat, more or less, about, almost, most, nearly, near, nigh, virtually, well-nigh, slenderly, quite, rather, to some extent, in some degree, partially, partly, part, after a fashion, well, sometimes, little, lightly, softly, less, a bit, a trifle, barely, just, pretty much, thinly
is (not)	When the sentence does not contain any degree measurement words.
not	hardly, barely, scarce, ill, never, non, no, none, least of all

3.3 Product affective analysis

People have different perceptions on linguistic variables. In order to deal with the uncertainty and approximation of the linguistics, this paper adopts fuzzy set theory. Fuzzy set theory has been applied in many different fields. It assumes that people use fuzzy sets to think rather than using precise numbers (Zadeh, 1965). A fuzzy set is defined by a crisp set and a membership function. The grade of membership of an element indicates the degree to which the element belongs to the fuzzy set. In this study, we use the triangular membership function that is commonly used in many studies (e.g. Yan & Nakamori, 2010; Chou, 2016). The fuzzy sets of linguistic degrees are shown in Figure 3. After the process of single-document affective analysis, a text is converted to a set of Kansei attributes and their corresponding linguistic degrees. The linguistic degrees are fuzzified as a fuzzy value and the degree of membership is weighted by the number of values of each Kansei attribute. Then, the clipped membership functions of the fuzzy values of the same Kansei attribute of each product are aggregated into a single fuzzy set. The fuzzy set is then defuzzified by the center of gravity (COG) method (Hirota et al., 1998). The aggregated results of the consumer reviews of the product represent the consumer affective response, and the aggregated results of the product description of the product represent the manufacturer's affective intention. The results of the product are visualized based on graphical representation of the affective profile. The following is an example of the product affective analysis process:

Given a set of pairs of a Kansei attribute and its corresponding linguistic degree extracted from consumers' reviews of a product:

{comfortable, is}, {comfortable, very}, {comfortable, very}

Fuzzification:

{comfortable, (0.333 is)}, {comfortable, (0.667 very)}

Fuzzy aggregation:

The aggregation of the fuzzy sets is show in Figure 3.

Defuzzification:

The COG value is 0.830

Given a set of pairs of a Kansei attribute and its corresponding linguistic degree extracted from a product description of the product:

{comfortable, extremely}

Fuzzification:

{comfortable, (1.0 extremely)}

Defuzzification:

The COG value is 0.958

Output affective profile:

Assuming that the degrees of other attributes are equal, the affective profile is shown in Figure 4.

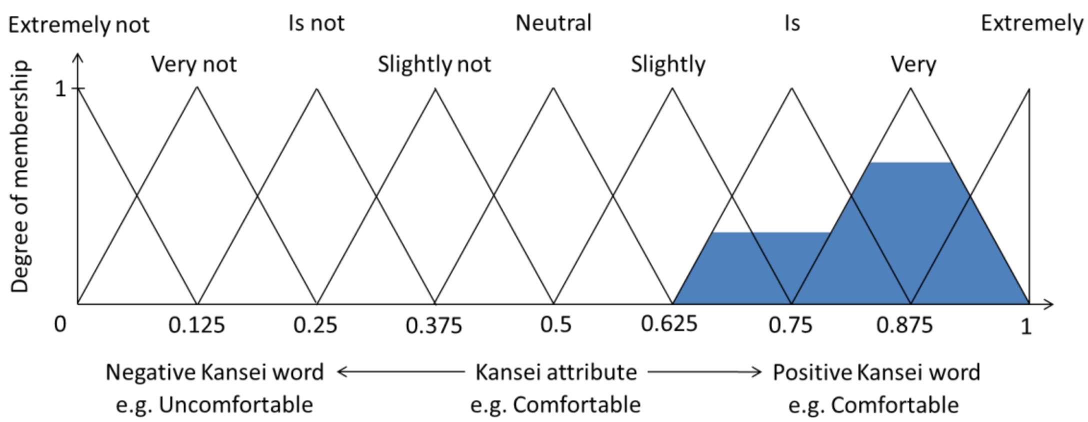


Figure 3 Membership functions of a Kansei attribute and aggregation of fuzzy sets

Figure 4 An example of affective profile

3.4 Applications

Affective information extracted from unstructured text can have different applications in affective design. In this paper, we focus on the gap analysis and product recommendations. We assume that the manufacturer's affective intentions of a product can be reflected in the use of wordings of the product description. And we also assume that the consumer's affective response is reflected in the use of wordings in the consumers' reviews. Thus, the information of the affective profile can be used to analyze the gap between the consumer's affective responses and manufacturer's affective intention.

On the other hand, affective information provides useful information for product recommendations. The affective recommender system (ARS) is a latest trending area of research, since publications in this area are few and recently published (Katarya & Verma, 2016). Traditional product recommendation systems use a variety of information, such as consumer demographics, consumer preferences, consumer behaviors, product features and keywords. The recommendation system has three commonly used methods, namely, content-based, demographic-based, and collaborative-based (Pazzani, 1999).

Content-based recommendations analyze product meta-data to calculate the similarities between products. It recommends to consumers about products that are similar to those purchased by consumers before. Keywords are often used to describe the product. The similarities are measured based on keyword matching. For example, Narducci et al. (2016) recommend products using keywords in different languages, Albatayneh et al. (2014) used latent semantic analysis to measure the similarity between product keywords.

Demographic-based recommendations measure the consumer similarity based on demographic information (e.g. gender, age, education level, etc.). Products are recommended to consumers based on products purchased by other consumers with similar demographic information. The recent application of this method can be found in Zhao et al. (2014), using demographic information extracted from social media for product recommendation.

Collaborative-based recommendations measure the similarity of consumers based on the purchasing model. Then, the system will recommend the products purchased by the similar consumers to the consumer. Recent developments of collaborative recommendation include Li et al. (2017), Pan and Ming (2016), etc. Collaboration via content recommendation is similar to Content-based recommendation. In contrast, collaboration via content recommendation measures product similarity based on the number of common consumers of the products. This method is widely used in Amazon and Alibaba, as well as different online shopping portals. The pages of these portals always provide information about the products that other consumers also brought and/or viewed.

In this study, we use affective information for product recommendation. The following sections describe the details of the assessment.

4. Case study

In order to demonstrate the usefulness of the proposed method, a case study is presented. We used a dataset provided by Julian McAuley on his web page[1]. He used the data in his study of collaborative filtering recommendation system (He and Mcauley, 2016). The dataset contains product reviews and metadata from Amazon, including 142.8 million reviews across 24 different product categories from May 1996 to July 2014. A product review is composed of a reviewer ID, a product ID, and a review text. A product metadata includes a product ID (asin), a product title, a description, a price, a brand name, categories, and a list of links of related products. The links of related product include the product IDs of the “also bought” products, and the “also viewed” products.

Following is a sample review based on the JSON format:


```
{
  "reviewerID": "AVQDEM0DQJC0B",
  "asin": "B000LQB24G",
```


[1] http://jmcauley.ucsd.edu/data/amazon/

"reviewerName": "Amazon Customer \"pawnman\"",
"reviewText": "This is a handy item to use…",
"unixReviewTime": 1372809600,
"reviewTime": "07 3, 2013"
}

Following is a sample metadata based on the JSON format:

{
'asin': 'B000LQB24G',
'title': 'Metro Vacuum SK-1 Air Force Blaster Sidekick Compact & Portable Motorcycle Dryer',
'categories': [['Automotive', 'Motorcycle & Powersports', 'Parts']],
'description': 'Thoroughly dry your bike or sports car in just eight …',
'price': 70.46,
'brand': 'Metro Vacuum',
'related':
{
'also_bought': ['B000WK4EC8', 'B000N5OOQ8', 'B000WJX6IM'],
'also_viewed': ['B001J4ZOAW', 'B00ABYVTXM', 'B0000CCXWA']
}
}

In this paper, we used a subset of the data of the automotive category, which each product should consist of a product description, at least one consumer review, at least one "also bought" product, and at least one "also viewed" product. There are 16,090 reviews, including 1490 products and 2917 consumers. Some statistics of the dataset are shown in Table 3.

Table 3 Statistics of the dataset

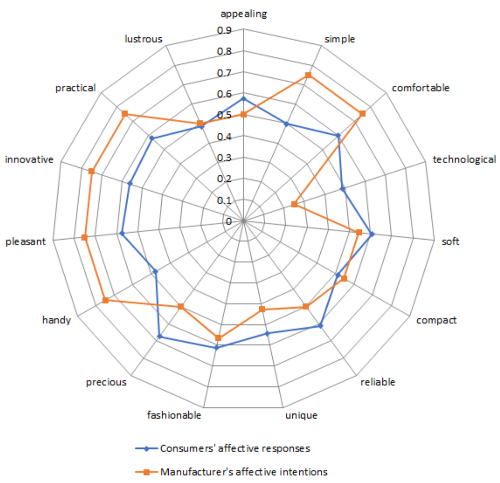
	No of texts	Min	Max	Average	Standard deviation
No. of words per consumer review	16090	1	2239	78.22	88.97
No. of words per product description	1490	2	516	75.44	62.85
No. of consumer reviews per product	1490	4	169	10.80	11.03
No. of consumer reviews per consumer	2917	1	42	5.51	3.00
No. of words of reviews per product	1490	105	13860	844.72	1024.35
No. of words of reviews per consumer	2917	20	1193	107.36	90.99
No. of Kansei words of reviews per product	1490	5	701	47.81	55.63
No. of Kansei words of product description	1490	0	29	4.69	4.33
No. of Kansei words of reviews per consumer	2917	0	63	7.33	5.22

4.1 Affective profile

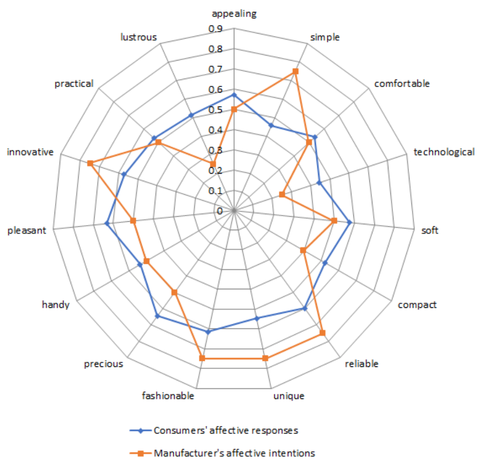
Based on the method described in the previous section, the Kansei attributes and the Kansei words are collected. We build an affective profile of each product. Five products are then selected as examples of the gap analysis to investigate the differences between the manufacturer's affective intention and the consumer's affective response. The five products include the most reviewed product (most review), the product with the longest reviews (longest review), the product with the longest product description (longest description), the product with the largest number of Kansei words in the reviews (most Kansei review), and the product with the largest number of Kansei words in the product description (most Kansei description). Table 4 shows the product ID, number of consumer reviews, number of words of product description and consumer reviews, and number of Kansei words of product description and consumer reviews. Figures 5(a) to 5(e) show the affective profile of the selected products. We also aggregate all the affective information of all products to investigate the overall affective profile of the data, as shown in Figure 5(f).

Table 4 Information of the selected products for gap analysis

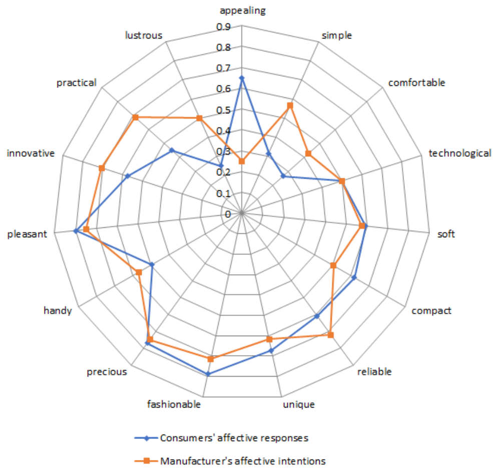
	Most review	**Longest review**	**Longest description**	**Most Kansei review**	**Most Kansei description**
Product ID	B000CITK8S	B00B7GC50Y	B000LQB24G	B007TG7HFO	B009XR48MM
No. of reviews	**169**	66	5	118	38
No. of words of reviews	12119	**13860**	249	11058	6845
No. of words of description	98	93	**516**	69	421
No. of Kansei words of reviews	569	673	23	**701**	389
No. of Kansei words of description	7	6	22	10	**29**



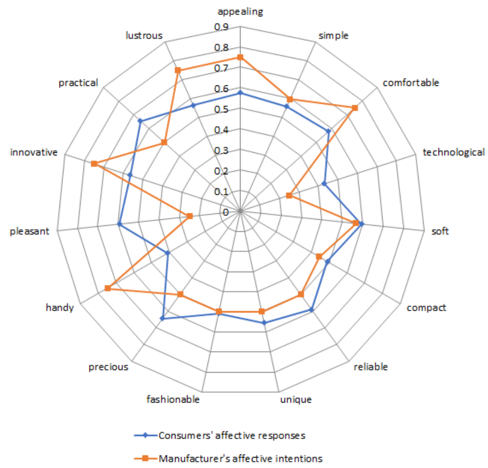
(a) B000CITK8S



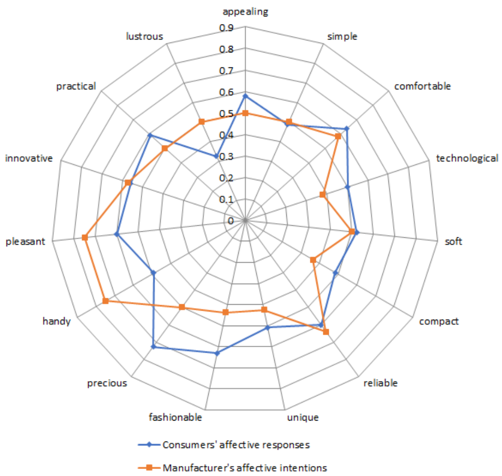
(b) B00B7GC50Y



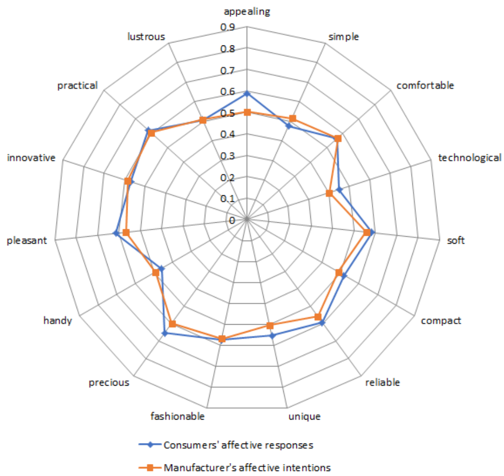
(c) B000LQB24G



(d) B007TG7HFO



(e) B009XR48MM



(f) All products

Figure 5 Affective profiles of the selected products

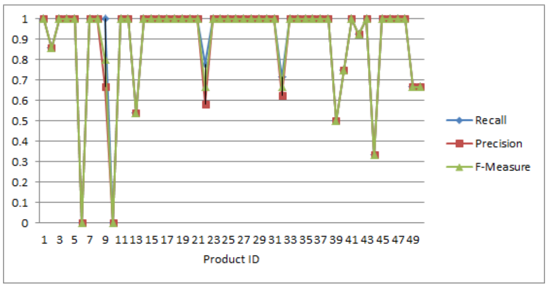
Based on the profile, we can easily observe the gap between the consumer's affective response and the manufacturer's affective intention. It is interesting to note that the affective attributes of the consumer affective responses are more neutral than that of the manufacturer affective intentions. This is because manufacturers always state the benefits of their products, while consumers might mention both advantages and disadvantages of the products. For example, in the product "B0000AY3X0", the product description is "The Absorber … could be one of the best drying tools you'll ever discover. The secret of the Absorber's power is it's amazing uniform… This design enhances capillary action and gives the Absorber it's super drying ability.".

On the other hand, we chose the automotive products as the dataset, but it is interesting that the technological attribute tends to be classical rather than high-tech. In general, most manufacturers always say that their products are comfortable, classical, soft, reliable, fashionable, precious, pleasant, innovative, and practical. Most consumers always point out that the products are appealing, comfortable, soft, reliable, unique, fashionable, precious, pleasant, innovative, and practical. The gap between the attributes of appealing, simple, technological, compact, reliable, unique, precious, handy and pleasant is relative large. In particular, manufacturers should pay more attention to the attributes that the degree of consumer response is lower than the manufacturer expectation, including simple, technological and compact.

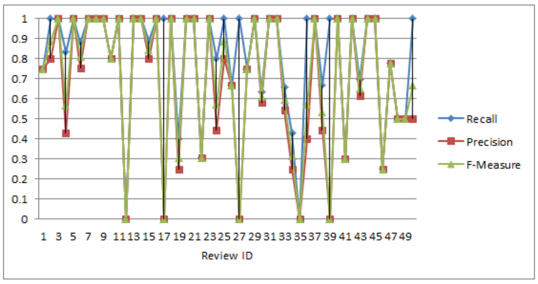
The accuracy of the extraction of the pairs of Kansei word and degree is accessed. 50 product descriptions and 50 reviews are selected based on the product ID. We manually construct the gold standard for the 100 text. The results are measured by recall, precision, and F-measure as shown in Figure 6 and Table 4. Based on the results, we can see that for both product descriptions and reviews, the averaged recall of affective information extraction is higher than 0.8. It indicates that the proposed method can effectively extract the affective information. The recall and precision of affective information extraction of product descriptions is higher than that of product reviews. This is because the product descriptions are relatively short and more straightforward. Consumers sometimes provide some suggestions that do not directly describe the products. Therefore, it is easier to extract the relevant information from the product descriptions.

Table 4 Evaluation results of extraction of pairs of Kansei word and degree

	Recall	Precision	F-Measure	Avg. no. of pairs
Product description	0.895	0.882	0.887	4.34
Product review	0.810	0.684	0.708	6.60



(a) Product description



(b) Product review

Figure 6 Evaluation results of extraction of pairs of Kansei word and degree

4.2 Product recommendation

In addition to the affective gap analysis, we also evaluated the usefulness of affective information on product recommendation. 18 experiments were performed by using different types of data and analysis. Table 5 shows the input data, methods, objectives, gold standards, and measurement methods. The experiments can be divided into Type I, Type II and Type III. Type I methods are intended to suggest product's similar products. This is useful when a consumer views a product, some of the related products can be displayed as a recommendation. Type II methods are intended to suggest consumer's similar consumers and then use similar consumers' reviewed products as a recommendation for the consumer. It is useful for providing personalized recommendations for known consumers based on the information of the consumers. Type III methods are intended to suggest products by comparing the similarity between products and consumers. It provides personalized recommendations to provide products with affective information similar to the consumers' affective information.

We use the related products (i.e. also bought products and also viewed products) as the gold standard for Type I methods. And we use the last reviewed product as the gold standard for Type II and Type III methods. In other words, we use the previous reviews of a consumer to find his/her similar consumers, and then we obtain the products reviewed by the similar consumers as recommendations, and finally we use the latest reviewed product of the consumer to check the accuracy. As a result, when analyzing the Type II and Type III methods, the consumers who have only 1 review are deleted. The last review of each consumer is also not used as input data. Since Type II and Type III methods have only 1 answer, we only use recall to evaluate. For Type I methods, we use recall, precision, and F-measure. The top n similar products of each product of Type I and Type III methods are selected using the threshold n, where n = 1, 2, 3, … and 100. For Type II methods, the top n similar consumers of each consumers are selected using the threshold n, where n = 1, 2, 3, … and 20.

Table 5 Setup of the experiments

ID	Input data	Method	Objective	Threshold	Answer	Measurement
1	Product information (i.e. price, brand, and categories)	SWA	Type I	100	RP	R, P, F
2	Kansei words of product description of each product	JS	Type I	100	RP	R, P, F
3	Kansei words of product reviews of each product	JS	Type I	100	RP	R, P, F
4	Affective profile of product description of each product	SWA	Type I	100	RP	R, P, F
5	Affective profile of product reviews of each product	SWA	Type I	100	RP	R, P, F
6	Unigram of product description of each product	JS	Type I	100	RP	R, P, F
7	Bigram of product description of each product	JS	Type I	100	RP	R, P, F
8	Unigram of product reviews of each product	JS	Type I	100	RP	R, P, F
9	Bigram of product reviews of each product	JS	Type I	100	RP	R, P, F
10	Common reviewed products of consumers (excluding last reviewed product)	JS	Type II	20	LP	R
11	Kansei words of product descriptions of products that reviewed by each consumer (excluding last reviewed product)	JS	Type II	20	LP	R
12	Kansei words of product reviews of each consumer (excluding last review)	JS	Type II	20	LP	R

13	Affective profile of product descriptions of products that reviewed by each consumer (excluding last reviewed product)	SWA	Type II	20	LP	R
14	Affective profile of product reviews of each consumer (excluding last review)	SWA	Type II	20	LP	R
15	Kansei words of product description of each product and Kansei words of product reviews of each consumer (excluding last review)	JS	Type III	100	LP	R
16	Kansei words of product reviews of each product and Kansei words of product reviews of each consumer (excluding last review)	JS	Type III	100	LP	R
17	Affective profile of product description of each product and affective profile of product reviews of each consumer (excluding last review)	SWA	Type III	100	LP	R
18	Affective profile of product reviews of each product and affective profile of product reviews of each consumer (excluding last review)	SWA	Type III	100	LP	R

Keys for similarity measurement methods:
SWA: Simple weighted average
JS: Jaccard similarity
Keys for objectives of the analysis:
Type I: To find similar products of each product
Type II: To find similar consumers of each consumer
Type III: To find products that have similar affective pattern of each consumer
Keys for referenced model answers:
RP: Related products (i.e. also bought, also viewed)
LP: Last reviewed product of each consumer
Keys for measurement methods:
R: Recall
P: Precision
F: F-measure

In Type I methods, they are content-based recommendations. 9 different types of data are used, including product information (i.e. categories, price, and brand of the products); Kansei words, affective profile, unigram, and bigram of the product descriptions; and Kansei words, affective profile, unigram, and bigram of the product reviews. In particular, n-gram is widely used in natural language processing. It is a common method of approximate matching in the fields of computational linguistics (Xia et al., 2011). N-gram is a continuous sequence of n words for a given text.

Unigram is an n-gram of size 1, and bigram is an n-gram of size 2.

For the method of using product information (Method 1), we use a simple weighted method to obtain the product similarity. This study used equal weights. The calculation is shown in Equations (1) to (4).

$$Sim_{method-1}(p_i, p_j) = \frac{Sim_{category}(p_i, p_j) + Sim_{price}(p_i, p_j) + Sim_{brand}(p_i, p_j)}{3} \quad (1)$$

$$Sim_{category}(p_i, p_j) = \frac{C_{p_i} \cap C_{p_j}}{C_{p_i} \cup C_{p_j}} \quad (2)$$

Where:

C_{p_i} and C_{p_j} are the sets of categories of products p_i and p_j respectively

$$Sim_{price}(p_i, p_j) = \frac{\min(d_{p_i}, d_{p_j})}{\max(d_{p_i}, d_{p_j})} \quad (3)$$

Where:

d_{p_i} and d_{p_j} are the price of products p_i and p_j respectively

$$Sim_{brand}(p_i, p_j) = \begin{cases} 1 & if\ b_{p_i} = b_{p_j} \\ 0 & otherwise \end{cases} \quad (4)$$

Where:

b_{p_i} and b_{p_j} are the brand name of products p_i and p_j respectively

For the methods of using Kansei words, unigram, and bigram of the product descriptions and the product reviews (Method 2, 3, 6 to 9), each product is indexed by its Kansei words, unigram, or bigram. Product similarity is then obtained based on Jaccard similarity method. The calculation is shown in Equation (5).

$$Sim_{method-n}(p_i, p_j) = \frac{W_{p_i} \cap W_{p_j}}{W_{p_i} \cup W_{p_j}} \quad (5)$$

Where:

W_{p_i} and W_{p_j} are the sets of Kansei words, unigram, or bigram of reviews or product description of products p_i and p_j respectively, and n = 2, 3, 6, 7, 8, 9

Each product is also represented by its affective profile (Methods 4 and 5) and is compared with other products according to weighted average method. This study used equal weights. The similarity between products is obtained by Equation (6).

$$Sim_{method-n}(p_i, p_j) = \frac{\sum_k^m \left| a_{p_i}^k - a_{p_j}^k \right|}{m} \quad (6)$$

Where:

$a_{p_i}^k$ and $a_{p_j}^k$ are the value of Kansei attribute k of consumer responses of products (or manufacturing intention of products) p_i and p_j respectively, m is the total number of the Kansei attributes, and n = 4, 5

In Type II methods, they are intended to suggest similar consumers to consumers. We use the common review products as a kind of collaborative-based recommendations (Method 10). The similarity between the consumers is calculated by Equation (7).

$$Sim_{method-10}(r_i, r_j) = \frac{P_{r_i} \cap P_{r_j}}{P_{r_i} \cup P_{r_j}} \quad (7)$$

Where:

P_{r_i} and P_{r_j} are the sets of products that have been reviewed by both consumers r_i and r_j respectively

In Type II methods, we also use the affective information of product description and consumer reviews to provide demographic-based recommendations (Methods 11 to 14). Each consumer is indexed with the Kansei words of their reviews, the affective profile of their reviews, the Kansei words of descriptions of their reviewed products, and the affective profile of descriptions of their reviewed products. The similarity of consumers is obtained by the Jaccard similarity method for Methods 11 and 12, as shown in Equation (8). The similarity of consumers is obtained by the weighted average method for Methods 13 and 14, as shown in Equation (9).

$$Sim_{method-n}(r_i, r_j) = \frac{K_{r_i} \cap K_{r_j}}{K_{r_i} \cup K_{r_j}} \quad (8)$$

Where:

K_{r_i} and K_{r_j} are the sets of Kansei words of reviews of consumers (or descriptions of the products that were reviewed by the consumers) r_i and r_j respectively, and n = 11, 12

$$Sim_{method-n}(r_i, r_j) = \frac{\sum_k^m \left| a_{r_i}^k - a_{r_j}^k \right|}{m} \quad (9)$$

Where:

$a_{r_i}^k$ and $a_{r_j}^k$ are the value of Kansei attribute k of consumer responses of consumers (or manufacturing intention of the products that were reviewed by the consumers) r_i and r_j respectively, and m is the total number of the Kansei attributes, and n = 13, 14

In Type III methods, they suggest products based on the similarity of the affective information between the products and the consumers. We use Kansei words and affective profile of each product's description to represent product affective information, and we use the Kansei words and affective profile of product reviews of each consumer to represent consumer affective information. Then we calculate the similarity between the information. The formulas are shown in Equations (10) and (11).

$$Sim_{method-n}(r_i, p_j) = \frac{K_{r_i} \cap K_{p_j}}{K_{r_i} \cup K_{p_j}} \tag{8}$$

Where:

$K_{r_i}$is the sets of Kansei words of reviews of consumer r_i, K_{p_j} is the set of Kansei words of product description of product p_j, and n = 15, 16

$$Sim_{method-n}(r_i, p_j) = \frac{\sum_k^m \left|a_{r_i}^k - a_{p_j}^k\right|}{m} \tag{9}$$

Where:

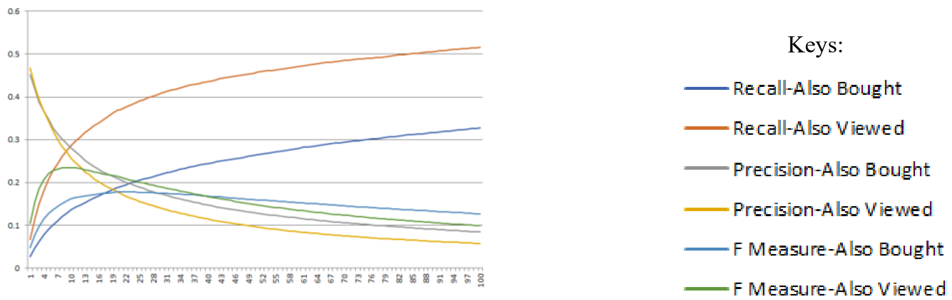
$a_{r_i}^k$ is the value of Kansei attribute k of consumer responses of consumer r_i, $a_{p_j}^k$ is the value of Kansei attribute k of manufacturing intention of the product p_j, and n = 17, 18

The evaluation results are shown in Figure 6 and Figure 7. The Y-axis is the recall, precision and F-measure, and the X-axis is the threshold. In Type I methods, it is apparent that the recall of the also viewed products is higher than that of the also bought products, and the precision of the also viewed products is lower than that of the also bought products. This is reasonable because people tend to view similar products rather than buy similar products. It is worth noting that when the threshold is small, the F-measure of the also viewed products is higher. When the threshold is large, the F-measure of the also bought products is higher.

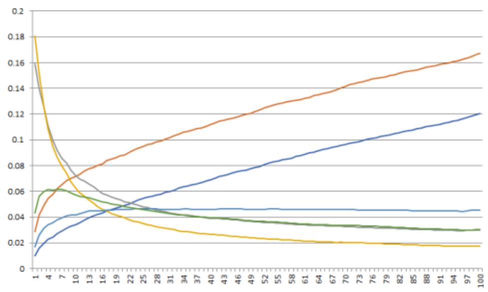
Among the Type I methods, the use of product information (i.e. price, brand, and categories) obtains the best result (Method 1). It shows that product information is useful for product recommendations. The results of using N-gram of product reviews (Methods 8 and 9) are better than that of product description (Methods 6 and 7). The results of using unigram (Methods 6 and 8) are better than that of using bigram (Methods 7 and 9). It shows that the bag of words of product reviews is useful for product recommendations. Among the use of affective information for product recommendation, the use of product description (Methods 2 and 4) is better than that of product reviews (Method 3 and 5) when the threshold is small. The use of product reviews is better than that of product description when the threshold is large. It shows that both information is useful for product recommendations. Using Kansei words of products (Methods 2 and 3) is always better than using affective profile of products (Method 4 and 5). This is because the affective profiles of products are highly abstracted, providing little information to provide good product recommendations.

Among the Type II methods, the best results are obtained by the use of information about common reviewed products (Method 10). The second one is the use of Kansei words of product descriptions of products that reviewed by each consumer (Method 11). The third one is the use of affective profile of product descriptions of products that reviewed by each consumer (Method 13). Using Kansei words and affective profile of product reviews of each consumer have the worst performance (Methods 12 and 14). It shows that the affective information of the product that reviewed by a consumer is more useful than the affective information of the product reviews of the consumer.

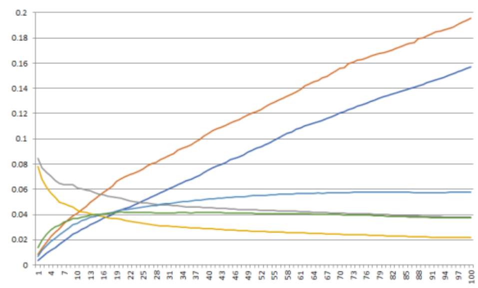
Among the Type III methods, the use of similarity between Kansei words of product description of each product and Kansei words of product reviews of each consumer (Method 15) is better than the other methods (Methods 16 to 18). Moreover, it obtains a relative high recall when the threshold is small (about 20). It shows that it is helpful using the affective information of product reviews to describe affective pattern of a consumer, and using the affective information of product description to describe affective pattern of a product.



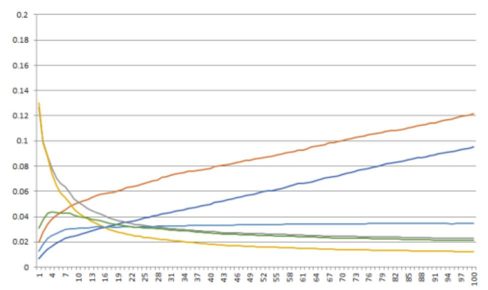
(a) Method 1



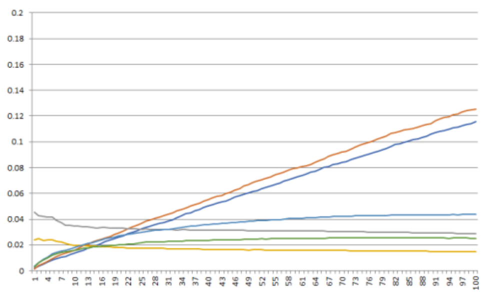
(b) Method 2



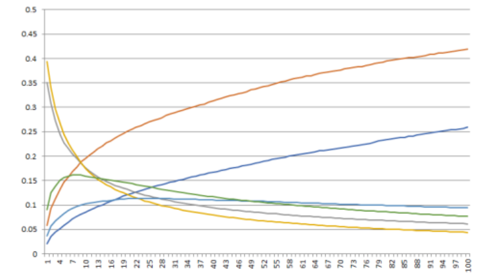
(c) Method 3



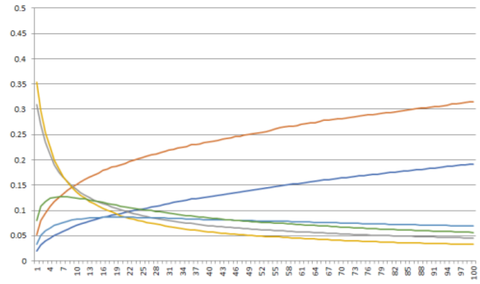
(d) Method 4



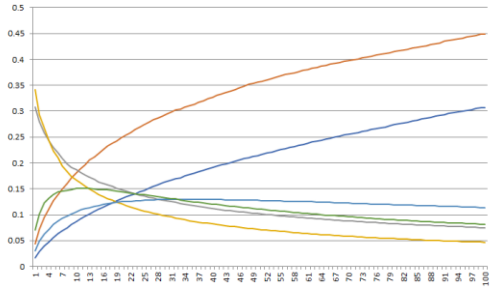
(e) Method 5



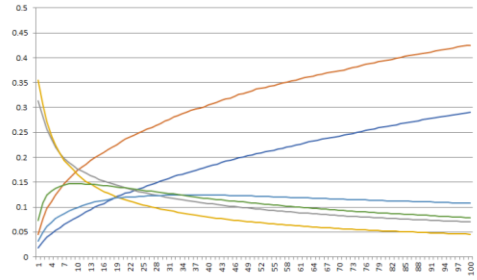
(f) Method 6



(g) Method 7

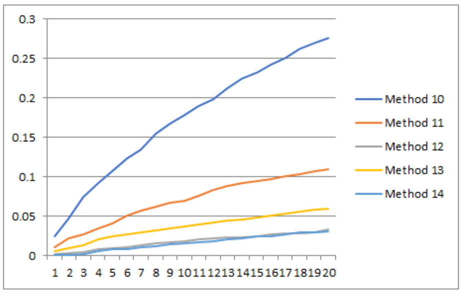


(h) Method 8

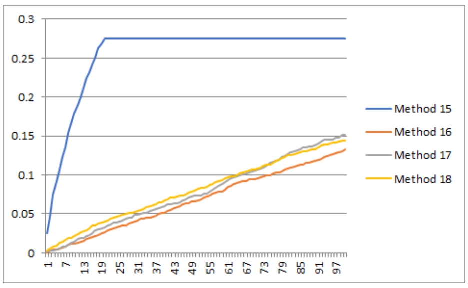


(i) Method 9

Figure 6 Evaluation results of Type I methods (X-axis is the threshold; Y-axis is the recall, precision and F-measure)



(a) Type II methods



(b) Type III methods

Figure 7 Evaluation results of Type II and Type III methods

5. Conclusion

The purpose of this paper is to develop an automated method of extracting the affective information of a product from the unstructured text at the early stages of affective engineering. We collect 15 common Kansei attributes from the literature. We develop a semi-automatic method to collect the relevant Kansei words. By using the collected Kansei attributes and Kansei words, we propose a text mining method that automatically extracts affective information from the product description and consumer reviews that reflect the manufacture's affective intention and the consumer's affective responses to the product. We assign a degree to each Kansei attribute of the product based on the fuzzy set theory. Therefore, we can build an affective profile for each product, and we can visualize the gap between the consumer's affective response and the manufacturer's affective intentions. To demonstrate the usefulness of the proposed method, we used Amazon.com's dataset to conduct a case study. We constructed the affective profiles of selected products and analyzed the gap between the manufacture's affective intention and the consumer's affective responses of the products. This information may help manufacturers to improve their products and services based on affective aspects. We also introduced the use of affective information in product recommendations. The results show that affective information helps to describe the affective patterns of products and consumers, and also contribute to product recommendations.

Suggested further work is as follows. 1) Product reviews always provide a rating to evaluate the usefulness of comments and products. We propose to use this information to further improve the accuracy of the Kansei attribute's degree analysis. 2) The use of affective information in product recommendation is relatively simple in this paper. It is suggested to combine different input data and methods to improve the accuracy. 3) The computational time of the proposed method increases linearly with the amount of data. Thus, it can be evaluated by using more data and applying it to other products or services. 4) In this study, we focus on the automatic construction of product and consumer affective profiles. We suggest using text mining method to discover the relationship between design elements and affective attributes, which will contribute to affective design.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (51405089), and the Science and Technology Planning Project of Guangdong Province (2015B010131008).

References

Albatayneh, N. A., Ghauth, K. I., & Chua, F. F. (2014). A Semantic Content-Based Forum Recommender System Architecture Based on Content-Based Filtering and Latent Semantic Analysis. Recent Advances on Soft Computing and Data Mining. Springer International Publishing.

Bahn, S., Lee, C., Chang, S. N., & Yun, M. H. (2009). Incorporating affective customer needs for luxuriousness into product design attributes. Human Factors and Ergonomics in Manufacturing & Service Industries, 19(2), 105-127.

Barone, S., Lombardo, A. & Tarantino, P. (2007) A weighted logistic regression for conjoint analysis and Kansei engineering, Quality & Reliability Engineering International 23.6(2007):689–706.

Boran, F. E., Efe, B., Akay, D., & Henson, B. (2014). Understanding customers' affective needs with linguistic summarization. KEER 2014 - International Conference on Kansei Engineering and Emotion Research.

Chan, K. Y., Kwong, C. K., Dillon, T. S., & Fung, K. Y. (2011). An intelligent fuzzy regression approach for affective product design that captures nonlinearity and fuzziness. Journal of Engineering Design, 22(8), 523-542.

Chou, J. R. (2016). A kansei evaluation approach based on the technique of computing with words. Advanced Engineering Informatics, 30(1), 1-15.

Fellbaum, C. (1998). WordNet. Blackwell Publishing Ltd.

Friborg, O., Martinussen, M., & Rosenvinge, J. H. (2006). Likert-based vs. semantic differential-based scorings of positive psychological constructs: a psychometric comparison of two versions of a scale measuring resilience. Personality &

Individual Differences, 40(5), 873-884.

Fung, C. K. Y., Kwong, C. K., Chan, K. Y., & Jiang, H. (2014). A guided search genetic algorithm using mined rules for optimal affective product design. Engineering Optimization, 46(8), 1094-1108.

Grimsæth, K. (2005) Kansei Engineering: Linking Emotions and Product Features, Norwegian University of Science and Technology, Report, Norwegian

Guo, F., Liu, W. L., Cao, Y., Liu, F. T., & Li, M. L. (2016). Optimization design of a webpage based on Kansei Engineering. Human Factors and Ergonomics in Manufacturing & Service Industries, 26(1), 110-126.

He, R., & Mcauley, J. (2016). Ups and Downs: Modeling the Visual Evolution of Fashion Trends with One-Class Collaborative Filtering. International Conference on World Wide Web (pp.507-517). International World Wide Web Conferences Steering Committee.

Hirota, K., Yoshino, H., Xu, M. Q., Zhu, Y., Li, X. Y., & Horie, D., (1998). A fuzzy case based reasoning system for the legal inference. In Fuzzy systems proceedings, IEEE world congress on computational intelligence, the 1998 IEEE international conference (Vol. 2, 4–9) (pp. 1350–1354), May 1998 .

Hsiao, Y. H., Chen, M. C., & Liao, W. C. (2017). Logistics service design for cross-border e-commerce using kansei engineering with text-mining-based online content analysis. Telematics & Informatics. 34, 284-302

Hsu, S. H., Chuang, M. C. , & Chang, C. C. (2000). A semantic differential study of designers' and users' product form perception. International Journal of Industrial Ergonomics, 25(4), 375-391.

Huang, Y., Chen, C. H., & Khoo, L. P. (2012). Kansei clustering for emotional design using a combined design structure matrix. International Journal of Industrial Ergonomics, 42(5), 416-427.

Jiang, H., Kwong, C. K., Liu, Y., & Ip, W. H. (2015a). A methodology of integrating affective design with defining engineering specifications for product design. International Journal of Production Research, 53(8), 2472-2488.

Jiang, H., Kwong, C. K., Siu, K. W., & Liu, Y. (2015b). Rough set and PSO-based ANFIS approaches to modeling customer satisfaction for affective product design. Advanced Engineering Informatics, 29(3), 727-738.

Jiao, J., Zhang, Y., & Helander, M. (2006). A kansei mining system for affective design. Expert Systems with Applications, 30(4), 658-673.

Katarya, R., & Verma, O. P. (2016). Recent developments in affective recommender systems. Physica A: Statistical Mechanics and its Applications, 461, 182-190.

Kwong, C. K., Jiang, H., & Luo, X. G. (2016). AI-based methodology of integrating affective design, engineering, and marketing for defining design specifications of new products. Engineering Applications of Artificial Intelligence, 47, 49-60.

Li, G., Zhang, Z., Wang, L., Chen, Q., & Pan, J. (2017). One-class collaborative filtering based on rating prediction and ranking prediction. Knowledge-Based Systems.

Li, Ming, and H. B. Yan. (2016) Applying Kansei Engineering to service design: A case study of budget hotel service. International Conference on Service Systems and Service Management IEEE, 2016.

Llinares, C., & Page, A. F. (2011). Kano's model in kansei engineering to evaluate subjective real estate consumer preferences. International Journal of Industrial Ergonomics, 41(3), 233-246.

Nagamachi, M., & Lokman, A. M. (2016). Innovations of Kansei engineering. CRC Press.

Nagamachi, M., 1989. Kansei engineering approach to automotive. J. Soc. Autom. Eng. Jpn. 43 (1), 94–100.

Narducci, F., Basile, P., Musto, C., Lops, P., Caputo, A., Gemmis, M. D., Iaquinta, L., & Semeraro G. (2016). Concept-based item representations for a cross-lingual content-based recommendation process. Information Sciences, 374, 15-31.

Osgood, C.E., Suci, G., & Tannenbaum, P. (1957). The measurement of meaning. Urbana, IL: University of Illinois Press.

Pan, W., & Ming, Z. (2016). Collaborative recommendation with multiclass preference context. IEEE Intelligent Systems, 32(2), 45-51.

Pazzani, M. J. (1999). A Framework for Collaborative, Content-Based and Demographic Filtering. Kluwer Academic Publishers.

Rosler, F., Battenberg, G., Schüttler, F. (2009). Subjective perceptions and objective characteristics of control elements. ATZautotechnology 9, 48-53.

Shieh, M. D., Yeh, Y. E., & Huang, C. L. (2016). Eliciting design knowledge from affective responses using rough sets and kansei engineering system. Journal of Ambient Intelligence and Humanized Computing, 7(1), 107-120.

Vieira, J., Osório, J. M. A., Mouta, S., Delgado, P., Portinha, A., Meireles, J. F., & Santos, J. A. (2017). Kansei engineering as a tool for the design of in-vehicle rubber keypads. Applied Ergonomics, 61, 1-11.

Xia, R., Zong, C., & Li, S. (2011). Ensemble of feature sets and classification algorithms for sentiment classification. Information Sciences, 181(6), 1138-1152.

Yan, H. B., & Nakamori, Y. (2010). A probabilistic approach to Kansei Profile generation in Kansei engineering. IEEE International Conference on Systems Man and Cybernetics (pp.776-782). IEEE Xplore.

Yan, H. B., Huynh, V. N., Murai, T., & Nakamori, Y. (2008). Kansei evaluation based on prioritized multi-attribute fuzzy target-oriented decision analysis. Information Sciences, 178(21), 4080-4093.

Zadeh, L. A. (1965). Fuzzy sets. Information and control, 8(3), 338-353.

Zhao, W. X., Li, S., He, Y., Wang, L., Wen, J. R., & Li, X. (2015). Exploring demographic information in social media for product recommendation. Knowledge & Information Systems, 49(1), 61-89.