Journal of Engineering Design

ISSN: 0954-4828 (Print) 1466-1837 (Online) Journal homepage: http://www.tandfonline.com/loi/cjen20

A product affective properties identification approach based on web mining in a crowdsourcing environment

Danni Chang & Carman Lee

To cite this article: Danni Chang & Carman Lee (2018): A product affective properties identification approach based on web mining in a crowdsourcing environment, Journal of Engineering Design, DOI: 10.1080/09544828.2018.1463514

To link to this article: https://doi.org/10.1080/09544828.2018.1463514

Published online: 18 Apr 2018.

JOURNAL OF ENGINEERING DESIGN, 2018
https://doi.org/10.1080/09544828.2018.1463514

Taylor & Francis
Taylor & Francis Group

A product affective properties identification approach based on web mining in a crowdsourcing environment

Danni Chang [a] and Carman Lee [b]

[a]School of Design, Shanghai Jiao Tong University, Shanghai, People's Republic of China; [b]Department of Industrial and Systems Engineering, Hong Kong Polytechnic University, Hong Kong, People's Republic of China

ABSTRACT

Affective product design, which aims to satisfy customer feelings as an aspect of product quality, has attracted more and more research attention. However, conventional product design relies on surveys and user experiments to collect user evaluations, which leads to the issues that (i) consumers can only express their feelings on design attributes specified by assigners; (ii) abundant online consumer resources are neglected; and (iii) a lack of further prioritisation and re-construction of affective design properties. This study aims to develop a product affective properties identification approach. Crowdsourcing platforms have the advantage of obtaining large numbers of free consumer comments and have been utilised as data sources. Web mining and text mining are deployed to capture the crowdsourced product review resources. Then product design knowledge hierarchy is established to find design properties, while sentiment analysis was undertaken to identify affections. With the help of domain ontology to connect design properties and corresponding affections, product affective properties can be identified. Furthermore, the identified affective properties are prioritised, so as to assist in design improvement and support decision making. To illustrate the proposed approach, a pilot study on iPhone 7 was conducted, in which influential affective properties have been identified and ranked.

ARTICLE HISTORY
Received 27 September 2017
Accepted 8 April 2018

KEYWORDS
Product affective property; crowdsourcing; web mining; product design knowledge hierarchy; domain ontology

1. Introduction

Nowadays, the increasingly competitive market has elicited an urgent need to develop successful products that can satisfy the increasing consumer expectations and demands (Cross 2000). Apart from basic functions and economic considerations, affective aspects of products are also of great concern to consumers (Jordan 1998; Khalid 2006) Design attributes, such as colour and form can provoke feelings, and influence the overall perception of a product. Therefore, affective product design (APD) was advocated to develop products that satisfy customer feelings as an aspect of product quality (Nagamachi 2002; Lai, Chang, and Chang 2005) In this regard, in the relevant literature, it is indicated that products with deliberate affective design can help improve consumer satisfaction and further promote product

CONTACT Danni Chang dchang1@sjtu.edu.cn

success (Seva, Duh, and Helander 2007) Therefore, good affective features can sharpen the competitive edge of products, and a precise understanding of product affective properties appears particularly important and deserves in-depth investigation.

Regarding previous affective design studies, most research focus was placed on the analysis of emotions (Pengnate and Sarathy 2017; Gennaria et al. 2017; Huang, Chen, and Khoo 2012) and the establishment of quantitative models linking emotions and design attributes (Nagamachi 2002; Park and Han 2004; Zhai, Khoo, and Zhong 2009) In terms of data acquirement methods, surveys and user experiments are widely adopted. As in (Huang, Chen, and Khoo 2012; Huang, Chen, and Khoo 2012) questionnaires are used as the main method to collect consumer responses. However, the design of a survey, questionnaire or user experiment, to some extent, imposes the constraints of user involvement and user freedom in presenting their feelings on any design properties.

Moreover, with the rapid development of the Internet techniques, Web 2.0 enables stronger interaction and participation of Internet users and leads to the participatory web (Hedges and Dunn 2018; Newman et al. 2016) Recently, the debate about Web 3.0 brings about the idea of further merging intelligent webs and web services, and correspondingly, the advancement of intelligent media and social media mining has deepened the connections between web service and users (Newman et al. 2016) New collaboration forms can be facilitated to hear customers' voice and involve them in product design process (Djelassi and Decoopman 2013) Therefore, Web technologies can provide interactive platforms for enterprises to connect with worldwide consumers, so as to improve customer participation in the product development process and leverage open innovation. For consumers, they have more convenient channels in which to contribute their opinions. In particular, crowdsourcing, which is an important method for drawing on large numbers of people to contribute their opinions (Cross 2000; Howe 2006; Chang, Chen, and Lee 2014) has become an important way to effect consumer responses. Taking Proctor & Gamble as an example, the use of a crowdsourcing platform '*InnoCentive*' to collect product problems and possible solutions from Internet users has helped them increase the problem solving rate to 30%. For more examples, Wikipedia, Amazon's Mechanical Turk and iStockPhoto.com take advantage of the tremendous numbers of Web users that are willing to contribute their knowledge and ideas. Crowdsourcing appears to be a promising way to solicit consumer responses and is studied as the main data source in this work. Considering that large numbers of consumers' comments are collected via crowdsourcing and are often presented by textual data, data mining techniques, which are efficient in dealing with big data and effective for textual analysis are considered.

Therefore, to fully consider users' opinions and discover possible important affective design properties from abundant online consumer resources, this paper aims to investigate affective design properties based on users/consumer responses acquired by crowdsourcing platforms. Furthermore, the identified affective properties are re-organised into a systematic structure according to their design and affective importance in order to provide designers with strategic reference for further improvement.

To make the explanation clearer, the concepts of product affective properties and affective performance are defined.

Definition 1. Product Affective Properties. Design aspects or product features which can provoke users'/customers' emotions. If there exists a cause-and-effect relation (in a qualitative or quantitative manner) between design properties and consumers' affections, the design properties are identified as product affective properties.

Definition 2. Affective Performance of Product Properties. It relates to the success of product properties in influencing consumers' emotions and can be measured against standards such as polarity and intensity.

2. Related works

In this section, existing related works are analyzed from three aspects. Firstly, the concept of crowdsourcing is examined to reveal the current use of interactive Internet platforms by consumers in order to contribute their ideas and opinions on the product design process. As an efficient way to collect large amounts of consumer responses, crowdsourcing environment will be focused as the source of consumer responses in this work. Then, the research status of product affective design is thoroughly reviewed, and potential promising research directions can be identified accordingly. Finally, data mining techniques, which are powerful in dealing with a large amount of data, are considered in APD. Based on the analysis of existing studies, the challenges and possible contributions of this work are summarised and highlighted.

2.1. *A brief overview of crowdsourcing*

The term 'crowdsourcing' was first coined by Howe in the article 'The Rise of Crowdsourcing', in 2006 (Howe 2006) Basically, crowdsourcing can be schematically depicted as in Figure 1. The employer/assigner (right side) defines the task requirements and submits the tasks to a mediator, viz. the online crowdsourcing platform. Online workers/providers (left side) who are interested in this task can work on it and submit their solutions to the mediator after completion. These solutions are then forwarded to the employer who rewards the participants if their solutions are approved.

Conventionally, crowdsourcing practices rely on the power of the massive work force. Nowadays, the Internet is developing rapidly, and modern crowdsourcing has transferred mainly to the Internet. Especially in the product design realm, the dependence on consumers/users is increasing, from the traditional internal development to the current

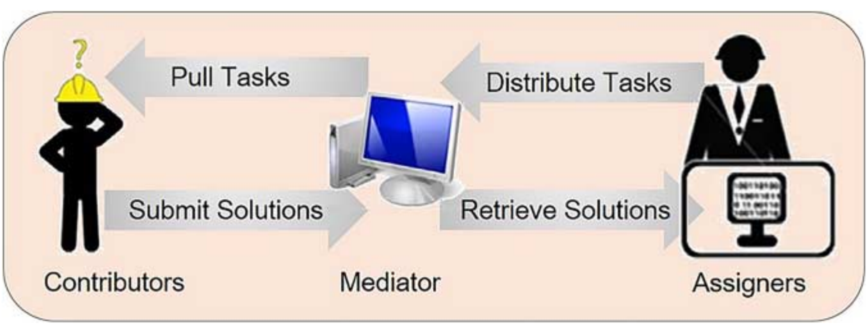

Figure 1. A typical crowdsourcing scheme.

user-centered design. It simultaneously requires more convenient channels for communication and interaction with customers/users. Compared with general websites, crowdsourcing platforms not merely enable Internet users to post their comments freely, but are also more targeted, having more Internet users with particular interest or knowledge of the product. Thus crowdsourcing may have higher possibilities to receive useful customer information. For this reason, crowdsourcing has been practiced in various design-related websites, such as Dell's *Ideastorm* for acquiring comments and suggestions for all Dell products from Internet users, product review websites gathering users' comments and idea generation websites collecting innovative design concepts. Therefore, crowdsourcing has assisted in developing a highly interactive environment for user/customer interaction and is an important source of consumer/user responses.

In product design and development practice, crowdsourcing can be realised through different activities to acquire user information. Based on the typical crowdsourcing scheme (as shown in Figure 1), specific tasks can be designed to suit different project purposes. Generally, the primary concern is the user involvement strategy. In this regard, participatory crowdsourcing (i.e. contributions from the users with interest) and opportunistic crowdsourcing (i.e. contributions from distributed sensors without users' active involvement) are two main strategies (Huang, Shema, and Xia 2017; Lane et al. 2010) In terms of the activities, crowdsourcing can be executed through contest (e.g. design competition, *DesignBoom*), collaboration (e.g. content co-creation, *Wikipedia*), collection (e.g. project-based ideation and collection, *iStockPhoto*) and labour (e.g. workforce employment, *Amazon MTurk*) (Boudreau and Lakhani 2013; Bal et al. 2017; de Mattos, Kissimoto, and Laurindo 2018) Due to the restriction of online platform, the common types of collected user information are text, image, audio and video. Nowadays, diverse crowdsourcing platforms have been introduced. For example, a requirement engineering crowdsourcing platform *CrowdREquire* has been developed to assist in the requirement elicitation, analysis and management (Adepetu et al. 2012) For idea generation, *OpenIdeo* and *Innocentive* are two important examples. Regarding crowdsourcing especially for design, there are *CrowdSpring*, *99Design*, *Cad Crowd* and *Designcrowd*. For crowdsourcing platforms which offer general crowdsourcing project development service, *Amazon Mechanical Turk* is one major example. Recently, the concept of crowdsourcing in the cloud is introduced, which aims at larger-scale and deeper crowd collaboration and development (Zhang et al. 2017) In this respect, *Amazon's AWS*, *Google Cloud Platform* and *Microsoft Azure* are three commonly used cloud platforms for many enterprises, and these platforms are beginning to show their power in supporting crowdsourcing. For example, AWS has become mature in supporting crowdsourcing service on MTurk. Therefore, the integration of crowdsourcing and cloud computing might be a promising direction of future crowdsourcing development.

In this study, the consumers' review and experience are target data. Based on the common approaches for consumers to offer their comments, the open platforms which employ participatory crowdsourcing to get consumers' text information is emphasised. The websites, which match crowdsourcing characteristics, can be considered as the data sources. To explain, the websites should be able to post a topic (i.e. the product or service) to attract Internet users with interest to contribute their ideas (i.e. comments), then the users with more high-quality participation will be rewarded. Therefore, product review websites, social media and product forums following the crowdsourcing scheme can be treated as the crowdsourcing data source.

In general, crowdsourcing platforms are advantageous in attracting users' contribution and have potential to get more valid user data, thus this study targets crowdsourcing websites as data source.

2.2. Affective product design

There have been many studies directed at APD, from different perspectives. Considering the basic assumption for APD that there exists a cause-and-effect relationship between consumers' affective responses and product features, existing studies can be generally classified into the following categories:

- *Identification and classification of consumers' affection*; this is to study consumers'/users' emotions provoked by products. Regarding particular techniques, the semantic differential (SD) method is frequently used to investigate customers' perception of products (Osgood, Suci, and Tannenbaum 1967) By studying product semantics, customers' subjective feelings about a product can be discerned and quantised on a Likert-type scale. Other assessment of emotions can be seen in the use of Conjoint analysis and Quality Function Deployment (QFD). Considering the ambiguity and subjectivity of consumers' affections, the selection and classification methods for Kansei words (or emotional adjectives) are a research hotspot. K-means clustering, affinity diagrams and design structure matrixes have been used to achieve this purpose (Huang, Chen, and Khoo 2012; Lokman and Kamaruddin 2010; Yang 2011)
- *Understanding of emotion-related design features*; for this research area, empirical study is a useful way and widely adopted by researchers. For example, an empirical survey was conducted to investigate the influence of website emotional design features, visual appeal and ease of use in regard to users' perceptions of usefulness, trust, as well as the intention to use websites (Pengnate and Sarathy 2017) An empirical experiment was performed to examine the emotions of children influenced by game design (Gennaria et al. 2017) In addition, quantitative Kansei models were developed to classify emotional features (Huang, Chen, and Khoo 2012) Moreover, the indexing of consumers' subjective and emotion-driven opinions is advocated to unify the 'perceived value' of designers and consumers from the affective perspective (Chen and Chu 2012)
- *Modelling the relationship between affective responses and design attributes*; for qualitative relations, hypothesis testing is often used to identify the effects of design factors on affective responses (Bhandari, Neben, and Chang 2017) However, most studies put research focus on the quantitative relations between the design attributes/factors and the affective responses. For example, a methodological framework combining user tests and statistical analysis was established to build links between user's emotional responses and the geometrical features of the products (Lu and Petiot 2014) A fuzzy regression model was proposed to model the relationship between customer satisfaction and product design parameters (Nazari-Shirkouhi and Keramati 2017) Moreover, Kansei engineering is a notable way of translating consumers' psychological feelings about a product into perceptual design attributes (Nagamachi 2002) A number of Kansei models have been accordingly developed to improve the association accuracy or applicability of Kansei approaches. To illustrate, fuzzy logic and rough set theory have been applied to cope with the uncertainty and ambiguity of affective responses (Park and Han

2004; Zhai, Khoo, and Zhong 2009) In addition, statistical methods, especially regression algorithms and artificial intelligence techniques, such as neural network, rule mining and genetic algorithms, are also widely studied to model the relationship between affective responses and specific design attributes (Hsiao and Tsai 2005; Fung et al. 2012) Recently emerging studies have placed more attention in tackling the nonlinearity problem of the relationships (Chan et al. 2011)

- *Design assessment and decision making*; this relates to the estimation of customer satisfaction based on affective responses, so as to support the specification, classification and prediction of the design attributes (Yang and Shieh 2010) Factor analysis and the hypothesis-testing approach are often used to examine the affective performance of the design attributes (Seva et al. 2011) Kansei engineering is also utilised to help with the classification of products for facilitating decision making in practical industrial design cases (Huang, Chen, and Khoo 2012) In this respect, pre-purchase affections are more emphasised in order to identify those affections which can influence the purchase decision (Seva, Duh, and Helander 2007; Seva, Duh, and Helander 2007)
- *Integration of affective design and other design concerns*; for instance, an artificial intelligence (AI)-based methodology for integrating affective design, engineering, and marketing for defining design specifications of new products has been proposed by which the concerns of the three processes can be considered simultaneously in the early design stage (Kwong, Jiang, and Luo 2016) In addition, (Seva, Duh, and Helander 2007) hypothesised that product attributes influence users' emotions and that the relationship is moderated by the adherence of these product attributes to purchase criteria, and further hypothesised that the emotional experience of the user influences purchase intention.

Although APD has been studied from different perspectives, the methods to acquire consumer data mainly rely on user tests including questionnaires, surveys and interviews. Online platforms, which have become a popular way for consumers to present their comments and thus contain considerable consumer resources, have been neglected. On the other hand, the design attributes used for such user tests are specified by assigners, rather than consumers. Specifically, product appearance such as colour and size is much more focused in APD (Chuang, Chang, and Hsu 2001; Crilly, Moultrie, and Clarkson 2004) In (Prado-León 2015) colour preference was studied in emotional design. (Seva et al. 2011) aimed to demonstrate that product attributes related to form are relevant in eliciting intense emotion and usability perception in mobile phones. (Zolkifly and Baharom 2016) intended to identify the roles of visual merchandising inside a car showroom as stimuli that attract customers. In addition, (Diego-Mas and Alcaide-Marzal 2016) presented a neural network based approach for modelling consumers' affective responses in product form design. Due to the major attention on product form design, other product features or design properties might have been neglected.

Furthermore, customers' affective evaluations elicited by product form directly lead to the issue that pre-purchase emotions are more focused. However, many essential product quality aspects, such as functionality, usability and safety, cannot be perceived unless customers use the product. Therefore, post-purchase or post-use affection is also important and deserves more research attention.

In summary, existing APD studies, in effect, impose constraints on consumers in expressing their feelings and opinions on any design attribute at any use phase using any preferable descriptions.

2.3. Data mining in affective product design

Intelligent techniques have been widely applied in APD. On the one hand, advanced quantitative models are developed to deal with the uncertainty and ambiguity in consumers' responses. For instance, rough set (RS) and particle swarm optimisation (PSO) based-ANFIS approaches were proposed to model customer satisfaction for affective design and to further improve the modelling accuracy (Jiang et al. 2015) A genetic algorithm (GA)-based rule-mining method was proposed to discover a set of rules relating design attributes with customer evaluation based on survey data in order to facilitate APD (Fung et al. 2012) On the other hand, intelligent computation is helpful for processing consumer data efficiently. Especially with the development of the Internet, the amount of consumer data involved in product design process is increasing massively.

Data mining, which is powerful in dealing with mass data, has attracted much attention. Actually, data mining is a generic term that covers such techniques as clustering (Lv, Lee, and Wu 2011; Sousa and Wallace 2006) association rule generation (Lee, Song, and Park 2012) and neural network (Cao, Li, and Ramani 2011; Chen and Liau 2001; Mazhar, Kara, and Kaebernick 2007; Park and Seo 2003) Clustering is mainly used for classification based on distances (similarities) between different concepts or designs. In APD, clustering is often applied to i) classify Kansei tags; and ii) classify product candidates (Huang, Chen, and Khoo 2012; Huang, Chen, and Khoo 2012) Association rule generation is employed to find regularities between products in large-scale transaction data. It is possible to cope with both numerical data and textual information. For example, rule-mining is utilised to discover a set of rules relating to design attributes based on customer evaluation so as to improve APD (Fung et al. 2012) Neural network consists of an interconnected group of artificial neurons, and processes information through a connectionist approach to computation. In this regard, a neuro network model was developed enabling single users' responses to different products to be predicted, and showed that the mathematical models based on neural networks achieved highly accurate predictions (Diego-Mas and Alcaide-Marzal 2016).

Besides, data mining is capable of coping with qualitative data, since online consumer responses are mostly in qualitative format (e.g. textual comments) (Yan et al. 2009). For example, web mining provides an effective way to discover textual patterns and to extract Web content; and text mining shows advantages in analyzing textual information and deriving high-quality information from the text. Regarding more specific applications, a framework was developed by He (2013) to improve user experience with case-based reasoning systems using text mining and Web 2.0. In addition, a text mining system, built on ontology to deal with the diagnosis data in the automotive domain, was proposed by Rajpathak (Rajpathak 2013) Moreover, text mining was applied in opinion polarity classification that helped decrease the sensitivity of ambiguous terms (Li and Tsai 2013).

In particular, Web Mining is advantageous at the discovery of the content and structure of Web resources (Stumme, Hotho, and Berendt 2006) For crowdsourced consumer reviews, the primary Web resources are individual pages containing consumer comments. In such context, Web content mining can be an effective way to retrieve online data, and can be used to group, categorise, analyze, and retrieve documents from crowdsourcing platforms. In this respect, Information Retrieval provides a range of popular and effective, mostly statistical methods for Web content mining. For example, swam intelligence has been used in document retrieval and clustering (Djenouri, Belhadi, and Belkebir 2018) Deep neural network has been used in web categorisation (López-Sánchez, Arrieta, and Corchado 2018) Evolutionary algorithm has been explored in assisting crowdsourcing quality control (Yung, Li, and Chang 2014) Association rule mining is often preferred in mining consumers' interests based on web database (Yeh et al. 2009).

Considering the nature of consumer comments, they are mostly semi-structured or unstructured textual data, thus text mining is commonly used jointly with web mining to further analyze Web content. Although online resources are tending to be multi-media, multimedia data mining widens the access to image, sound, video, and can produce semantic annotations to them. Therefore, text mining shows increasing power in dealing with online data. For example, Latent Semantic Analysis (LSA) has been used to realise webpage classification (Wang, Peng, and Liu 2015) Rough set theory has also been investigated to construct self-supervised text classification algorithm (Shi et al. 2011) Moreover, machine learning models are paid with much attention to extend text mining with adaptive computation capability (Cai et al. 2018) Convolutional neural network has been applied in opinion identified from text (Poria and Cambria 2016) In addition, ontology learning is widely discussed to improve the precision of the understanding of user comments (Maedche and Staab 2001) Accordingly, natural language processing, string matching and adaptive learning system have also been explored to enhance ontology-based text processing (Serra, Girardi, and Novais 2014).

Furthermore, as with the development of cloud computing and concepts such as Internet of Things and Cyber Physical System, the size of data is extremely expanded, and single machine or agent is not enough to deal with such data. Therefore, cloud environment and cloud platforms are introduced to realise the mining of larger-scale data and real-time collaboration of dynamic and elastic cloud resources (Amiri, Mohammad-Khanli, and Mirandola 2018; Tsai, Liu, and Wang 2017; Hwang et al. 2016) For new product design, the cloud platforms can enable customer resources integrated and utilised to profile more complete and precise user characteristics (Balco, Law, and Drahošová 2017).

Therefore, advances in information technology, data gathering and analytics, in effect, have enabled companies to manage all phases of the customer life cycle, including acquiring customer requirements, improving current design and service based on customer opinions and retaining customer data. Especially, web mining and text mining are suitable to cope with crowdsourced consumer opinions and will be considered in the proposed approach.

2.4. Summary

Based on a thorough analysis of existing related work, it was found that APD has become an important research area in the product design realm. However, challenges still exist, hindering the further exploitation of online consumer resources from the affective perspective. In detail, (1) Previous approaches commonly adopted surveys and questionnaires to collect consumers' affective responses, and neglected the massive online resources; moreover, the product attributes/design properties in survey are specified by assigners, which restricts users' choices to comment on other attributes; (2) Affective design is often considered for product form based on pre-purchase emotion, and the customer feedback covering the whole product life has not been fully investigated; (3) For existing product feature identification approach using text mining or ontology, the commonly adopted method is to incorporate expertise to pre-define product features, and then identify product features according to the predefined rules, that actually leads to the issue that the product features which can be identified are limited, and the information contained by the features which are not pre-defined will be lost; (4) Existing related research is heavily inclined to sentiment polarity identification, and the further analysis of product affective features (e.g. prioritising and re-constructing affective design properties) is still deficient.

To tackle the problems above, this work aims to contribute a product affective properties identification approach which is able to (1) make use of crowdsourced consumer responses (where free comments including pre- and post-use experience can be obtained) to identify important affective design properties; (2) deal efficiently with a large amount of data in qualitative format; and (3) further prioritise and represent affective design properties in terms of their comprehensive performance in both affective and design aspects.

Therefore, the novelty of this study mainly lies in (i) the utilisation of online crowdsourcing user resources to fully understand consumers' affective experience and capture their opinions covering the whole product life; (ii) the integration of design knowledge into product feature identification approach to achieve a complete and precise extraction of product attributes, and (iii) the prioritisation of product features based on both affective intensity and design importance to reach a comprehensive analysis integrating design and consumer concerns. Generally, this paper is intended to identify and prioritise product affective features based on crowdsourced user comments, so as to facilitate design practice with insights on the advantages and disadvantages of current design and indications about possible design improvement directions.

3. Research methodology

The overall framework of research methodology is outlined in Figure 2. Generally, it can be divided into three stages. Stage 1 is to capture product review data from crowdsourcing websites (i.e. online resources) and perform basic text processing to extract useful textual tokens. In Stage 2, the textual tokens are examined to identify product affective properties. For this purpose, a product design knowledge hierarchy (PDKH) is constructed to provide design considerations in order to identify design properties, and ontology is utilised to assist in semantic analysis and sentiment analysis, so as to investigate the emotions associated with the product properties. In Stage 3, a prioritisation process is deployed to estimate the importance of different affective design properties. As a result, affective design properties can be ranked according to their comprehensive importance in terms of design and consumers' affections, and a new representation of product properties can be mapped out which can be useful reference for designers in regard to design evaluation and decision making. Furthermore, the properties with strong affections will be retained for potential use in future design.

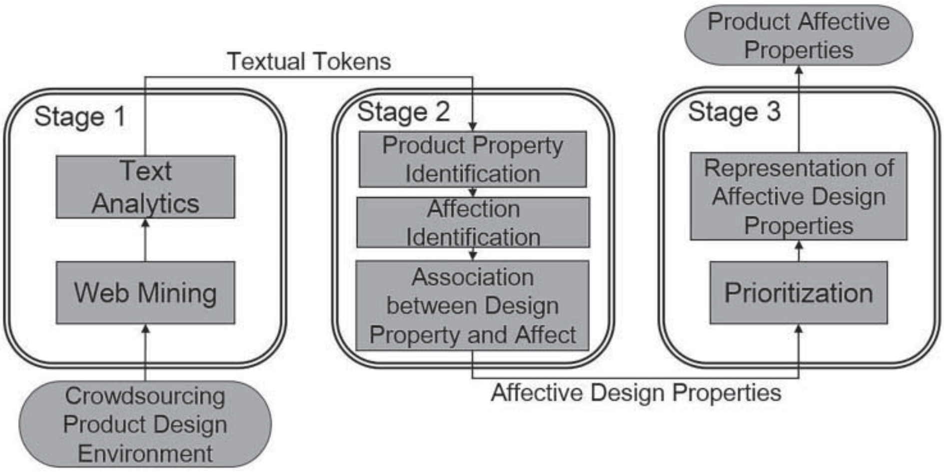

Figure 2. Overall framework of research methodology.

3.1. Stage 1: extraction of online consumer responses

As mentioned above, with the rapid development of the Internet, there are more and more online platforms and convenient access for consumers to offer their understanding and impressions of products. Customers/users can give any comments on any design aspects which can arose their interest. Especially for crowdsourcing, it creates channels to collect opinions and solutions from the large number of Internet users and is studied as an important data source of consumer responses. Therefore, the development or selection of a proper product design crowdsourcing platform should be the primary consideration. Actually, there are already some online platforms which can function as crowdsourcing platforms, such as Wikipedia. In regard to product design, pervasive sensor networks, internet services and social media, especially online product review websites, which actually contain abundant consumer responses, can provide a wide range of sources of customer responses. Therefore, existing crowdsourcing platforms, such as online product review websites, are treated as the main sources of consumer responses.

Step 1: ***Content Extraction***. The primary inputs are online crowdsourcing websites where consumers commonly post their comments and opinions. Web mining is applied to crawl web logs of crowdsourcing websites (e.g. product review websites) to extract meaningful content (i.e. consumer responses/comments). Considering the data formats contained in such websites are varied (e.g. textual data and graphics, which are mainly used in Web 2.0), textual data is mainly focused in this study, since product reviews are often expressed in user language, and posting texts is a comfortable and preferred way for consumers. Extracted content is collected as response documents. In particular, the content from one webpage is stored as one document, and all documents are treated as the corpus. Through this step, the web content, particularly the textual information of target web pages will be outputted.

Step 2: ***Text Processing***. Text mining is applied to deal with the captured contents. Necessary operations include the discovery of textual patterns, tokenization to divide the documents into textual tokens (including steps like setting filter stop words and filter tokens by length, transforming cases into a certain pattern, the nature is to separate the document into individual words) and further pre-processing of the textual data (e.g. calculating TF-IDF, correlation, and similarity in a quantitative manner). Considering the further processing for identifying the affective design properties, N-gram generation is employed to capture short phrases, so as to fully retain the semantic context and ensure the accurate identification of the affective product properties. For example, individual token 'headphone' is too generic, and any description including headphone will be counted. To improve the precision of feature extraction, n-gram terms can be defined to clarify the features to be extracted. Take the same example, 2-gram 'headphone jack' is more specific to pinpoint the physical port and functions related to the headphone jack, and 3-gram 'headphone lightening adapter' is more specific to represent the physical accessary to connect Earpod and the 3.5mm port.

Therefore, the inputs of Stage 1 are the crowdsourced consumer feedback and comments, and outputs are extracted individual tokens or n-gram short terms which are semantically meaningful (Figure 3).

3.2. *Stage 2: identification of product affective properties*

The extracted word tokens (including individual tokens and n-grams) are examined from the perspectives of design and affection, since the importance of product properties in design and consumer affections is not the same. For example, the cpu chip of a smartphone is very important from the design perspective; however, it is not a property directly attracting consumers' emotional responses. Therefore, design and affective aspects should be jointly considered to ensure the accurate capture of important emotional properties. In particular, tokens (or n-grams) representing product attributes or specifications are identified as design property tokens (or n-grams). Tokens implying affections, such as adjectives or nouns which have been defined as containing emotions in referring to some lexical database, are identified as affective tokens. Through identifying the semantic and design connections between affective tokens and design property tokens, product affective properties can be found.

Step 3: ***Establishment of PDKH***. From the design perspective, a product design knowledge hierarchy (PDKH) is outlined to assist in the identification of design-related tokens. With reference to (Garrett 2011) product design (as functionality) can be considered layer by layer from abstract ideas to concrete specifications (as shown in Figure 4). There are mainly five layers, i.e. product objectives, functional specifications, interaction design, interface design and sensory design. Following this structure, design can be developed in detail for each level. In particular, a hierarchical structure can be referred to for depicting the design knowledge at each level. As shown in Figure 5, design knowledge can be analyzed from: what aspects should be considered; what kind of properties should be equipped; to what specifications should be set at each design level. In general, Figure 4 presents the product design flow, while Figure 5 provides the knowledge representation structure for each design stage. With the help of Figures 4 and 5, design knowledge can be fully examined and systematically presented.

For different products, the details of PDKH should be specified accordingly. Then the tokens extracted from the crowdsourced responses can be analyzed with reference to

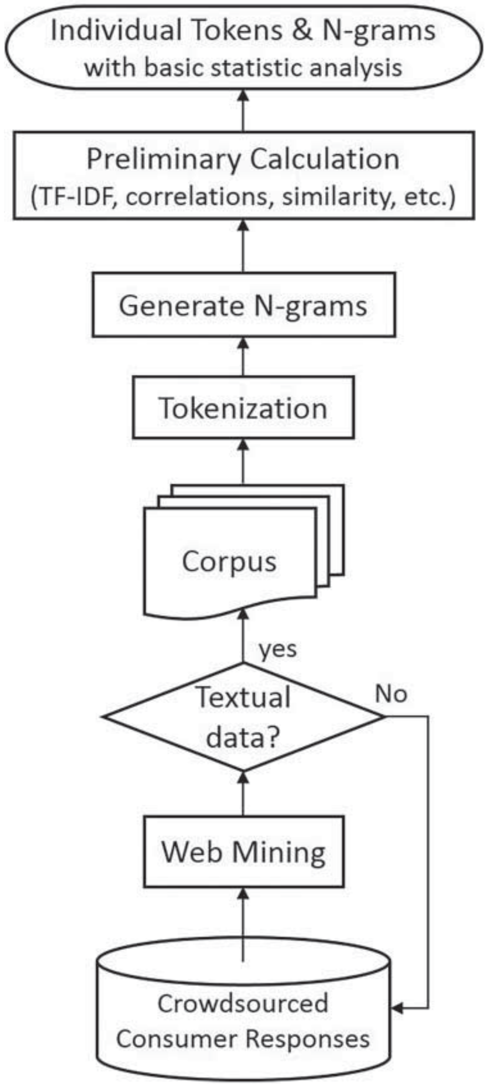

Figure 3. Workflow of Stage 1.

PDKH. For this purpose, a lexical database is needed to provide the Synset (semantic) relations and Word (lexical) relations. Particularly, a set of lexical relations can be adopted, including superordinate associations such as Hypernym (is a kind of), Holonym (is a part of), and subordinate association such as Hyponym (. . . is a kind of), Meronym (. . . is a part of) and Attribute (. . . is a value of). The extracted tokens will be analyzed referring to PDKH to identify design-related tokens. If tokens can be associated with the design knowledge in PDKH (e.g. subordination, similar, synonymous), the tokens can be identified as design tokens. Therefore, design properties can be found from each document.

Step 4: Affection Identification. For affective tokens, an electronic lexical database is necessary to provide the definition, lexical categories (e.g. nouns, verbs, adjectives and adverbs),

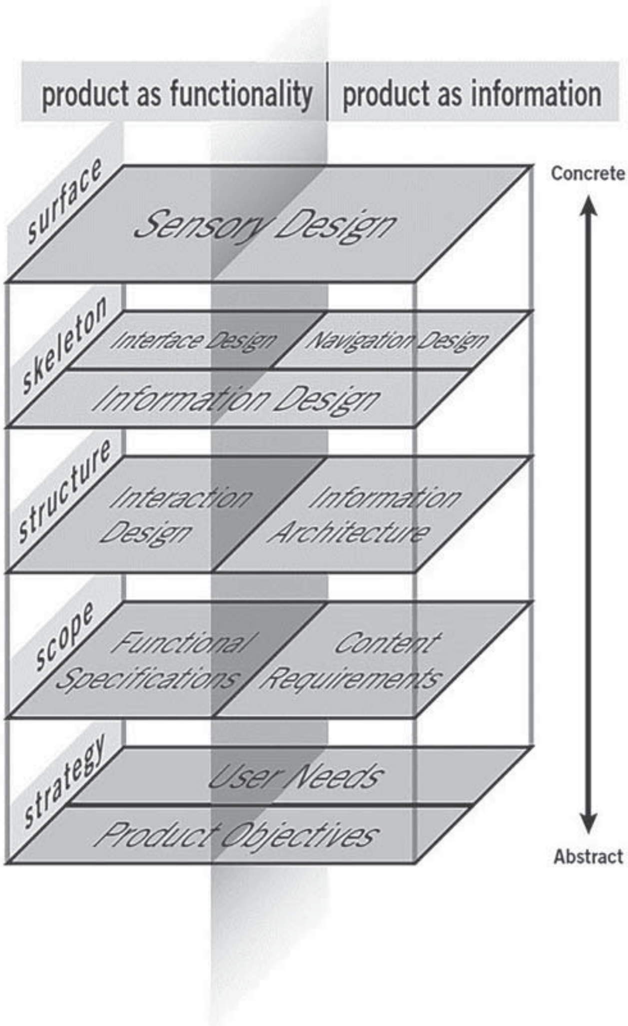

Figure 4. Five layers of design based on user experience (Garrett 2011).

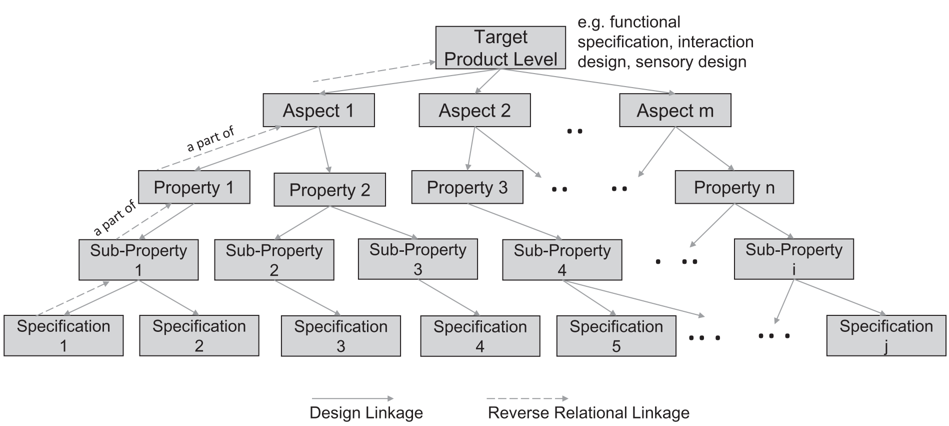

Figure 5. Product design knowledge hierarchy.

semantic relations and word relations to help with the identification of affection. Especially, a sentiment dictionary (e.g. The Dictionary of Affect in Language), which defines the sentiment contained by different words, is needed as important reference for polarity analysis and sentiment measurement. At this step, two main actions are executed, i.e. sentiment analysis to examine the affect (in the current study, polarity analysis is more empathised), and preliminary computation to outline the general sentiment trend.

Step 4.1: Sentiment analysis; In fact, emotions should be understood in context. The more fully the whole context is understood, the more exact is the estimation of the emotion or mood. Therefore, sentiment in this work is analyzed from two levels, i.e. sentence-level (local context) and document-level (whole context). The sentiment of certain words (or concepts in the lexical database) is estimated mainly in the local context, and whole context is considered to further amend the sentiment analysis results.

Firstly, polarity analysis is deployed to determine the sentiment polarity of each document. Polarity confidence can be calculated accordingly and treated as the sentiment score in whole context (denoted as SS_W and $SS_W \in [0, 1]$). With the help of the lexical database and sentiment dictionary, word tokens which contain affections can be identified. Particularly, existing sentiment dictionaries, which have been established based on consumer reviews (e.g. hotel, travelling), will be adopted as training data to support the identification of affect. Correspondingly, n-grams containing the affective tokens can be treated as affective n-grams. Afterwards, the sentiment of the sentences containing these affective tokens/n-grams can be analyzed, and the sentiment score in local context (denoted as SS_l and $SS_l \in [0, 1]$) can be obtained. Therefore, the sentiment score can be estimated by combining $SS_W$and SS_l. Regarding the calculation of sentiment intensity, it integrates the concerns about (i) qualifier, (ii) degree adverb, and (ii) negative adverb. Besides the sentiment dictionary for identifying qualifiers with consumer emotion, the rules defining the degree adverb and negative adverb need to be established. For example, if there is negative adverb like 'no', 'not' and 'do not', the sentiment polarity can be treated as opposite, and -1 will be multiplied. For degree adverbs like 'very' and 'extremely', different weights should be assigned.

If the polarity of the whole context is consistent with the local context, the sentiment is strengthened. Considering that the sentiment is measured mainly based on local context, the correction effect caused by the whole context can be treated as equivalent to degree adverbs, and thus a square operation can be applied (Negnevitsky 2011).

$$SS = SS_l \bullet (1 + |SS_W|^2)$$

If the polarity of the whole context is opposite to the local context, the sentiment is weakened.

$$SS = SS_l \bullet (1 - |SS_W|^2)$$

Step 4.2: Preliminary computation; Based on the sentiment analysis results in the previous step, basic statistical analysis can be executed to show the general quantitative trend. For example, the ratio of the positive and negative responses, the correlations and similarities between different responses or different tokens. Descriptive statistics can be considered to describe the quantitative characteristics of the collected response. In general, this step aims to depict the rough affective performance of the product.

Through sentiment analysis, tokens containing affections can be identified and treated as affective tokens.

Step 5: Establishment of Associations between Design Properties and Affections. In steps 3&4, product property tokens and affective tokens have been identified. The subsequent consideration is to identify the connections between them, so as to further identify the product affective property.

Generally, one complete consumer response is treated as one document, so the xth document can be denoted as D_x. The identified token of D_x is denoted as T_{xy}. Assume the total number of identified tokens of D_x is m. Then D_x can be denoted as:

$$D_x : \{T_{x1}, T_{x2}, T_{x3} \ldots T_{xm}\}$$

The corpus consists of all documents. Assume the total number of documents is n:

$$\phi : \{D_1, D_2, D_3, \cdots D_x \cdots D_n\} = \{(T_{11}, T_{12}, T_{13} \ldots T_{1m_1}), (T_{21}, T_{22}, T_{23} \ldots T_{2m_2}), \ldots \\ (T_{n1}, T_{n2}, T_{n3} \ldots T_{nm_n})\}$$

Domain ontology is leveraged to examine the semantic relations between affective tokens (AT) and product property tokens (DT). Generally, ontology consists of two main elements, i.e. concepts (entities such as shape, colour, material) and relations (such as IS-A, part-of). It is a formal and generic way to represent a set of related concepts, thus is useful to reveal the essential relationships between different entities. It has been widely used in the field of information systems, for example, constructing causal maps (Wang et al. 2008) supporting research management (Ma et al. 2012) enhancing adaptive learning (Lau et al. 2009) and understanding customer requirements and assisting collaborative design (Chen et al. 2013; Bock et al. 2010) In particular, domain ontology captures concepts and their relations in a specific domain and is able to represent the axioms (e.g. rules) and constraints that define the prominent features of the domain. In this study, the identification of product affective features relies on design knowledge and domain experience. Although data mining methods such as text mining can effectively extract and analyze texts or words, the inherent relations between different product concepts should be identified based on design principles and constraints. For this reason, design domain ontology is required to help identify the connections between different textual entities from design perspective. Therefore, domain ontology is adopted to identify product features.

Thus it is easy to show that the AT and DT of xth documents are both subsets of D_x:

$$AT_x \subset D_x; DT_x \subset D_x$$

A relationship set U is defined to bridge affective tokens $at_i \in AT$, $\mathrm{AT} = \{at_i | i = 1, \ldots, I\}$ and design property tokens $dt_j \in DT$, $\mathrm{DT} = \{dt_j | j = 1, \ldots, J\}$. As shown in Figure 6, the elements in AT are related to DT based on U, and an association matrix can be thusly obtained. The value of the elements of U is from the set {0, 1}. If AT and DT are connected with each other, the u is 1; if not, u is 0. If an association relationship exists in a proper semantic context, the dt with associated at can be regarded as an affective product property.

Moreover, the local context and n-grams are also referred to identify U. If AT and DT are in the same local context (i.e. sentence-level context) or the same n-grams, they can be treated as connected with each other.

In general, the inputs of Stage 2 are extracted tokens and n-gram tokens, and outputs are affect-feature-pair, namely, the product affective features (Figure 7).

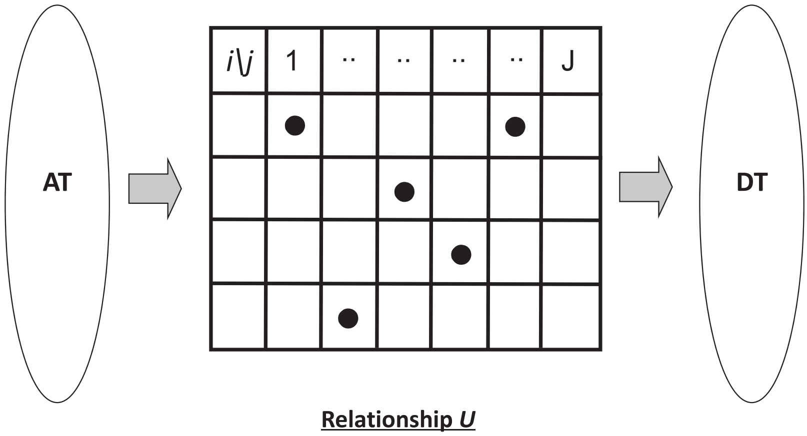

Figure 6. Association relationships between *AT* and *DT*.

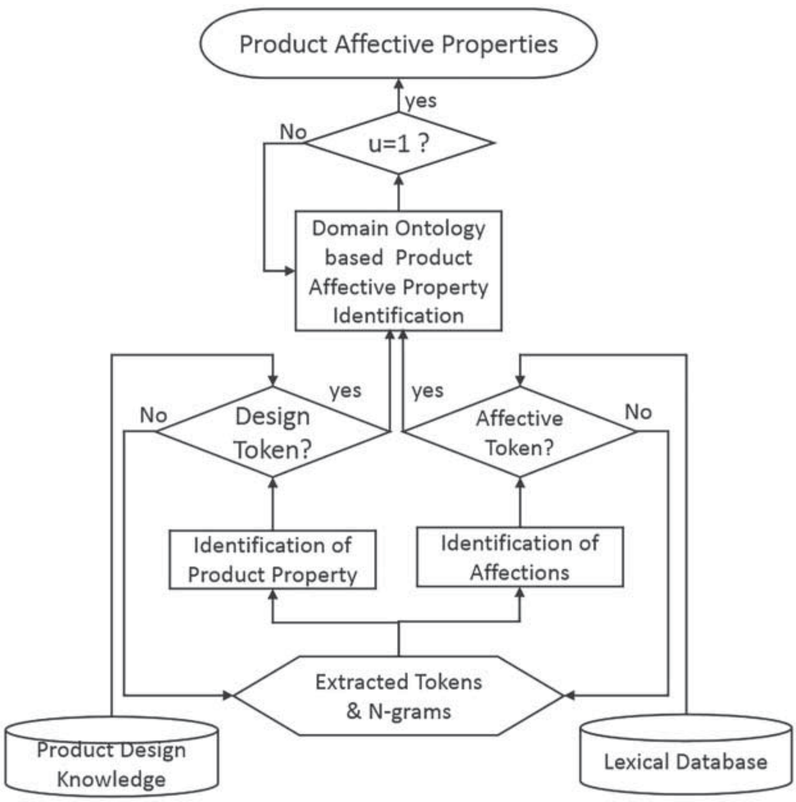

Figure 7. Workflow of Stage 2.

3.3. Stage 3: prioritization of product affective properties

The identified product affective properties are further prioritised according to their design importance and affective intensity. However, for the same design property, different consumers may have different emotional responses. Therefore, the effect on design properties is analyzed in their corresponding documents first, and then all the emotions on the same design properties are integrated to achieve the overall affective performance. Through combining the affective performance and design importance, the final priority of product affective properties can be obtained. Based on the priority, the product affective properties can be re-organised and ranked accordingly.

Step 6: Prioritisation. The priority of product affective properties is estimated from two perspectives: design and affect. The Design importance (*DI*) is calculated based on the total occurrence (i.e. $TO_{dt_{xi}}$, *i*th design token in *x*th document), term frequency–inverse document frequency ($TF-IDF_{dt_{xi}}$, see Equation 1) and *PDKH* priority (i.e. priority subject to *PDKH*, $f(h_{PDKH})_{dt_{xi}}$, see Equation 2) of the design property tokens. For $TO_{dt_{xi}}$ and $TF-IDF_{dt_{xi}}$, they are used to measure if the property dt_{xi} is of concern to consumers.

$$TF-IDF_{dt_{xi}} = tf(dt_{xi}, D_x) \cdot idf(dt_{xi}, \phi) = (0.5 + 0.5\frac{f_{dt_{xi},D_x}}{\max_{dt_{xi}} f_{dt_{xi},D_x}}) \cdot \log\frac{N}{n_{dt_{xi}}} \tag{1}$$

where f_{dt_{xi},D_x} is the number of times that design token dt_{xi} occurs in document D_x, N is total number of documents in the corpus ϕ, $n_{dt_{xi}}$ is the number of documents where the design token dt_{xi} appears.

According to (Hayes, Wheelwright, and Clark 1988) the earlier design stage is crucial to the final product quality and product life cycle cost, thus the design tokens at the more abstract levels should occupy heavier importance. With reference to (Ehrig et al. 2005; Song, Li, and Park 2009; Wang et al. 2008) the depth function is useful to assign importance to the hirarchical structure:

$$f(h_{PDKH})_{dt_{xi}} = \frac{e^{\beta h_{PDKH}} - e^{-\beta h_{PDKH}}}{e^{\beta h_{PDKH}} + e^{-\beta h_{PDKH}}}; \beta > 0 \tag{2}$$

where h_{PDKH} is the depth of the design token to the top level (i.e. the most concrete level) in PDKH, so the more abstract design leve dt_{xi} is located, $f(h_{PDKH})_{dt_{xi}}$ is heavier; β is a smoothing factor, $\beta > 0$.

The more often dt_{xi} is discussed, it indicates a greater importance to consumers. Moreover, if it only appears in one particular document, it implies that the design property is not of wide concern by consumers and may contain high risk of personal bias; thus, less importance should be assigned to this design property. Therefore, the design importance of dt_{xi}, i.e. $DI_{dt_{xi}}$ can be calculated using Equation 3.

$$DI_{dt_{xi}} = TO_{dt_{xi}} \cdot (1 - TF - IDF_{dt_{xi}}) \cdot f(h_{PDKH})_{dt_{xi}} \tag{3}$$

For the same design property dt_i, the design importance estimated in different documents can be integrated to achieve the overall design importance to the corpus:.

$$DI_{dt_i} = \sum_{x=1}^{N_{dt_i}} DI_{dt_{xi}} = \sum_{x=1}^{N_{dt_i}} TO_{dt_{xi}} \cdot (1 - TF - IDF_{dt_{xi}}) \cdot f(h_{PDKH})_{dt_{xi}} \tag{4}$$

where N_{dt_i} is the total number of documents containing the design property dt_i.

Likewise, the affective intensity is estimated based on the polarity (i.e.$p_{at_{xj}}$, if positive, then $p_{at_{xj}} = +1$; if negative, then $p_{at_{xj}} = -1$), the total occurrence of affective token (i.e.

$TO_{at_{xj}}$), $TF - IDF_{at_{xj}}$ (see Equation 5) and sentiment score $SS_{at_{xj}}$ (numerial value without polarity concern, see Equation 6).

$$TF - IDF_{at_{xj}} = tf(at_{xj}, D_x) \cdot idf(at_{xj}, \phi) = \left(0.5 + 0.5 \frac{f_{at_{xj},D_x}}{\max_{at_{xj}} f_{at_{xj},D_x}}\right) \cdot \log \frac{N}{n_{at_{xj}}} \tag{5}$$

where f_{at_{xj},D_x} is the number of times that affective token at_{xj} occurs in document D_x, N is total number of documents in the corpus, $n_{at_{xj}}$ is the number of documents where the design token at_{xj} appears.

As mentioned in Section 3.2, sentiment score is estimated based on the joint consideration of local and whole contexts in order to identify the exact emotional intensity of the affective token.

$$SS_{at_{xj}} = \begin{cases} SS_{lat_{xj}} \cdot (1 + |SS_{wat_{xj}}|^2), & if \quad p_{lat_{xj}} = p_{wat_{xj}} \\ SS_{lat_{xj}} \cdot (1 - |SS_{wat_{xj}}|^2), & if \quad p_{lat_{xj}} \neq p_{wat_{xj}} \end{cases} \tag{6}$$

where $SS_{lat_{xj}}$ is the sentiment score in local context (sentence-level), and $SS_{wat_{xj}}$is the sentiment score in the whole context (document-level).

Thus, the affective intensity of affective token at_{xj}, i.e. $AI_{at_{xj}}$, can be calculated using Equation 7.

$$AI_{at_{xj}} = p_{at_{xj}} \cdot TO_{at_{xj}} \cdot (1 - TF - IDF_{at_{xj}}) \cdot SS_{at_{xj}} \tag{7}$$

For the same design property dt_i, the associated affect in the whole corpus can be accumulated in two ways. One is the sum of absolute values of all associated affections to reflect the total intensity of the whole emotion attached to the property (see Equation 8). Another is the algebraic sum of all associated affections including their polarities in order to determine the overall polarity of the design property (see Equation 9).

$$|AI| \text{ for } dt_i = \sum_{x=1}^{N_{dt_i}} \sum_{j=1}^{n_{at_{xi}}} AI_{at_{xj}} = \sum_{x=1}^{N_{dt_i}} \sum_{j=1}^{n_{at_{xi}}} TO_{at_{xj}} \cdot (1 - TF - IDF_{at_{xj}}) \cdot SS_{at_{xj}} \tag{8}$$

where N_{dt_i} is the total number documents containing dt_i, $n_{at_{xi}}$ is the total number of affective tokens which are associated to design property dt_i in D_x.

$$AI \ for \ dt_i = \sum_{x=1}^{N_{dt_i}} \sum_{j=1}^{n_{at_{xi}}} AI_{at_{xj}} = \sum_{x=1}^{N_{dt_i}} \sum_{j=1}^{n_{at_{xi}}} p_{at_{xj}} \cdot TO_{at_{xj}} \cdot (1 - TF - IDF_{at_{xj}}) \cdot SS_{at_{xj}} \tag{9}$$

where N_{dt_i} is the total number documents containing dt_i, $n_{at_{xi}}$ is the total number of affective tokens which are associated to design property dt_i in D_x.

One product affective property represents one design property and the associated affectation. Thus an affective product property ($P_{xi,xj}$) can be denoted as $at_{xj} \cdot dt_{xi}$. A weighted calculation is introduced to integrate DI and AI in order to achieve the overall priority (i.e.

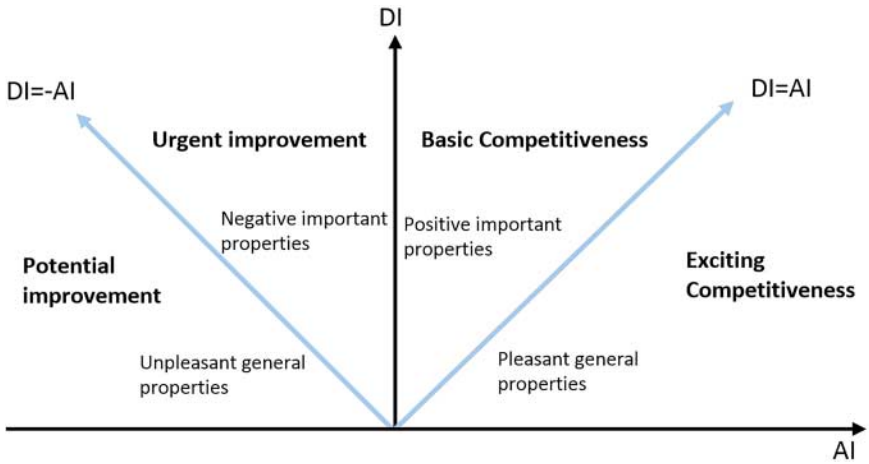

Figure 8. Strategic segmentation of product design properties according to *DI* and *AI*.

the priority of design property *i* with affect *j*, $OP_{xi,xj}$, Equation 10).

$$\begin{aligned} OP_{xi,xj} &= \omega_d DI_{dt_{xi}} + \omega_a |AI_{at_{xj}}| \\ \omega_d + \omega_a &= 1; 0 < \omega_d < 1 \& 0 < \omega_a < 1 \end{aligned} \tag{10}$$

where w_d is the weight of design importance, w_a is the weight of affective intensity; |AI| is considered, since no matter positive or negative responses, the more the absolute value, the more it is concerned by consumers.

For one product affective property, the overall priority (OP) to the corpus can be accumulated as:

$$OP_i = \sum_{x=1}^{N_{dt_i}} \sum_{j=1}^{n_{dt_i}} OP_{xi,xj} = \sum_{x=1}^{N_{dt_i}} \sum_{j=1}^{n_{dt_i}} (\omega_d DI_{dt_{xi}} + \omega_a |AI_{at_{xj}}|) \tag{11}$$

where N_{dt_i} is the total number documents containing dt_i, $n_{at_{xi}}$ is the total number of affective tokens which are associated to design property dt_i in D_x.

Step 7: Product Affective Property Representation. Every product affective property can be denoted as 'design property (*DI*, *AI*, *OP*)'. Generally, a higher *OP* means a higher integrated priority, thus more attention should be paid. In particular, for $AI > 0$, it represents the product properties with a positive consumer affection; therefore, such properties could be the strength of this design. On the contrary, for $AI < 0$, it indicates the product properties cannot satisfy the consumers and maybe a weakness in the product. As proposed in Figure 8, normalised *DI* and *AI* can be used to segment product properties into different sections according to their affective performance. Strength and weakness can be easily recognised, and valuable references as to which properties, to what degree, should be improved can be achieved.

In general, the inputs of Stage 3 are identified product affective features, and outputs are the prioritised product affective features with specific design importance and affective intensity.

4. A pilot study

A pilot study on iPhone 7 was conducted to demonstrate the proposed approach, since mobile phone design is often used for case studies in APD research. To facilitate the comparison between the proposed approach and other methods, the smartphone is used in the pilot study. As the development of crowdsourcing platform is not the focus of this study, existing crowdsourcing websites are considered as data sources. Amongst them, Amazon and CNET, which are two important product review platforms that often post topics for collecting comments from product users, are selected as the crowdsourcing product review resources. Therefore, review posts and comments under the topic of iPhone 7 are targeted.

Normally, the comments are presented in a combination of textual descriptions and product pictures. Considering that texts are widely preferred by Internet users to express their opinions, and image processing is beyond the scope of this work, textual content is the focus and non-textual content is removed. Moreover, considering that the qualities of the posted reviews vary greatly (viz. some are very rough and grammatical mistakes frequently appear) and not all of them contain sufficient valuable information, a filtering process is deployed. Too rough responses (less than 50 words) are removed, and responses with too many grammatical mistakes are removed, as well. Hence, 57 documents were captured in total in this preliminary study. The post dates range from 2016 September to 2017 May. The complete content of one review is treated as one document, and the collection of all documents is treated as the corpus.

4.1. Content extraction and text analytics

The first stage is actually a data mining process, where web mining and text mining are applied to extract useful textual information from the crowdsourced responses. For this purpose, a powerful tool *RapidMiner* is used to crawl the product review webpages and perform basic text analysis. The main processes are deployed as *Read Excel*, *Get Pages*, *Data to Documents*, and *Process Documents* (Figure 9). In particular, the target web links are collected and the URLs are recorded in excel. The *Read Excel* operator is used to read excel and recognise the details in each cell. The *Get Pages* operator is functional in crawling web sites, extracting HTML content and verifying if the content value matches the expected value type. Each design document is regarded as one example set. *Data to Documents* is utilised to convert the example sets to an object collection where each row represents each example document. Having the organised object collection, the *Process Document* operator is responsible for specifically coping with the example sets and is executed by the sub-processes of *Tokenise*, *Transform Cases*, *Filter Stopwords (by English)*, *Filter Tokens (by length)* and *Generate n-Grams (terms)*. Therein, *Tokenise* is used to separate the textual content into individual word tokens. *Transform Cases* is to convert the word tokens into a consistent case and avoid repetitive count of the same words, so that the text analysis is not case-sensitive. *Filter Stopwords (by English)* and *Filter Tokens (by length)* are employed to control the tokenization process and restrict the representation of the tokens. For example, if the value of *Filter Tokens (by length)* is set as five, that means the words longer than five letters are represented as the first 5 letters. To fully and correctly understand consumers' emotions, their original expressions are extremely important and should be retained. Therefore,

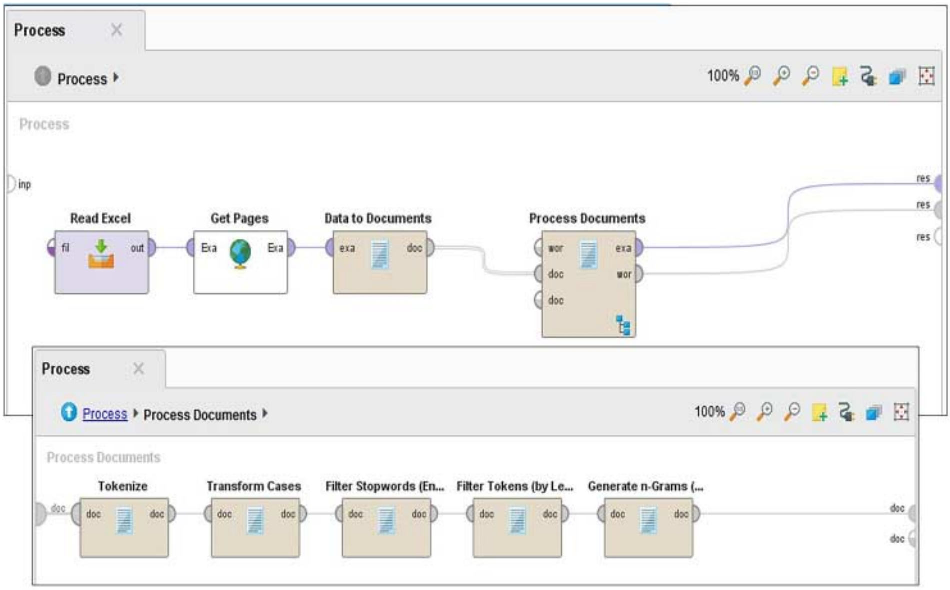

Figure 9. Extraction of online consumer responses with rapid miner.

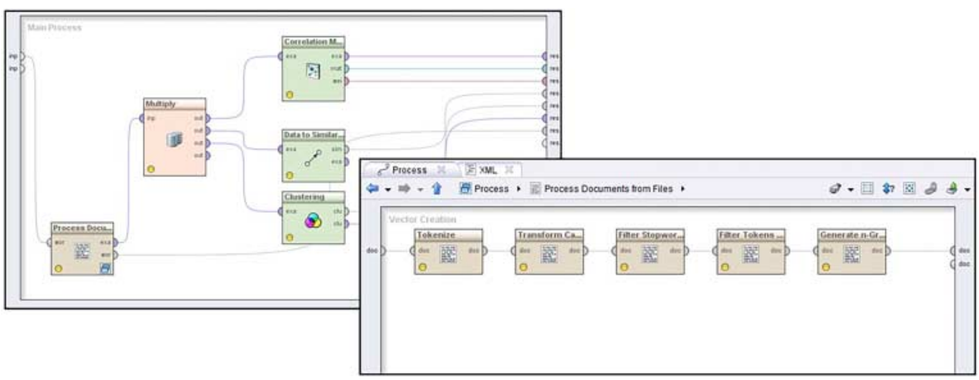

Figure 10. Preliminary analysis of online consumer responses with rapid miner.

Generaten-Grams (terms) (in this study, N is set as 3) are deployed to generate semantically meaningful short phrases.

In addition, *Correlation Matrix*, *Data to Similarity* and *Clustering* can also be considered to provide simple analysis of the captured documents (Figure 10). However, the correlations and similarities in these operators are calculated, mainly based on numerical vectors such as TF-IDF, rather than the affective performance.

14426 attributes (including single word tokens, 2-gram terms and 3-gram terms) were obtained (Figure 11).

The table below presents a list of examples of the extracted tokens (including single tokens, 2-gram tokens and 3-gram tokens) (Table 1).

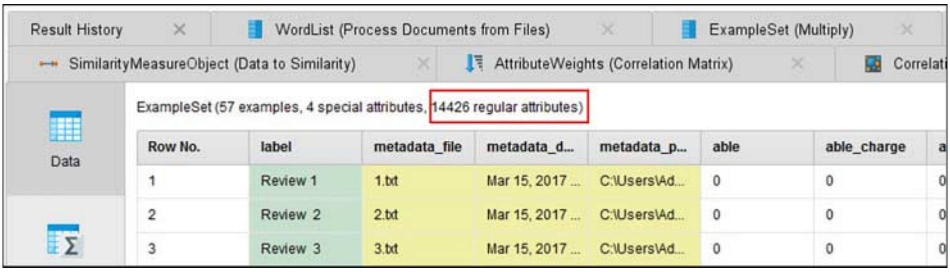

Figure 11. *RapidMiner* extraction results.

Table 1. Examples of extracted tokens.

absurd	loss	love
absurd_mounts	loss_headphone	love_android
absurd_mounts_cash	loss_headphone_jack	love_android_good
accept	loss_price	love_apple
accept_card	loss_price_pods	love_apple_products
accept_card_calls	lost	love_button
accept_card_phone	lost_android	love_button_took
accept_want	lost_android_root	love_fastest
accept_want_updated	lost_destroyed	love_fastest_device
access	lost_destroyed_expensive	love_improvements
access_apps	lost_easier	love_improvements_came
access_apps_basic	lost_easier_send	love_iphone
access_framerate	lost_focus	love_iphone_reason
access_framerate_resolution	lost_focus_beautiful	love_larger
access_tech	lost_time	love_larger_screen
access_tech_relied	lost_time_included	love_nonsense
accessories	lost_truth	love_nonsense_people
accessories_adaptors	lost_truth_matter	love_portrait
accessories_adaptors_integrated	Lots	love_portrait_mode
accessories_companies	lots_peoplelots_people_care	love_reviewlove_review_totally
accessories_companies_iphones edt.	etd.	love_tech

4.2. Identification of product affective properties

For the identification of product property tokens, PDKH is referred in examining if these word tokens are related to design knowledge of a Smartphone. An electronic lexical database *WordNet* is introduced to provide lexical and semantic references. By the use of *WordNet*, the definitions, lexical categories (e.g. nouns, verbs, adjectives and adverbs), semantic relations and word relations can be recognised.

Generally, the extracted word tokens are marked as *Concepts*, *Entities* and *Others*. The tokens tagged as *Concepts* and *Entities* are examined to determine if they belong to Smartphone design knowledge. A set of proper semantic relations is selected based on domain ontology to narrow the connections down to certain relations with relatively higher importance. Referring to previous ontology-related studies, '*synonymous*', '*Meronym*' (a part of) and '*Hypernym*' (a kind of) are frequently studied, and thus used in this study to identify the associations between the word tokens and PDKH. Two main kinds of associations are considered: (1) whether these tokens belong to any corresponding design levels, and (2) whether these tokens are semantically related to the tokens which have been confirmed to be design property tokens. To explain, Figure 12 shows the association process of tokens extracted from Document 1 based on PDKH. The PDKH needs to be constructed based on design knowledge and domain experience and relies on experts' manual effort. With the

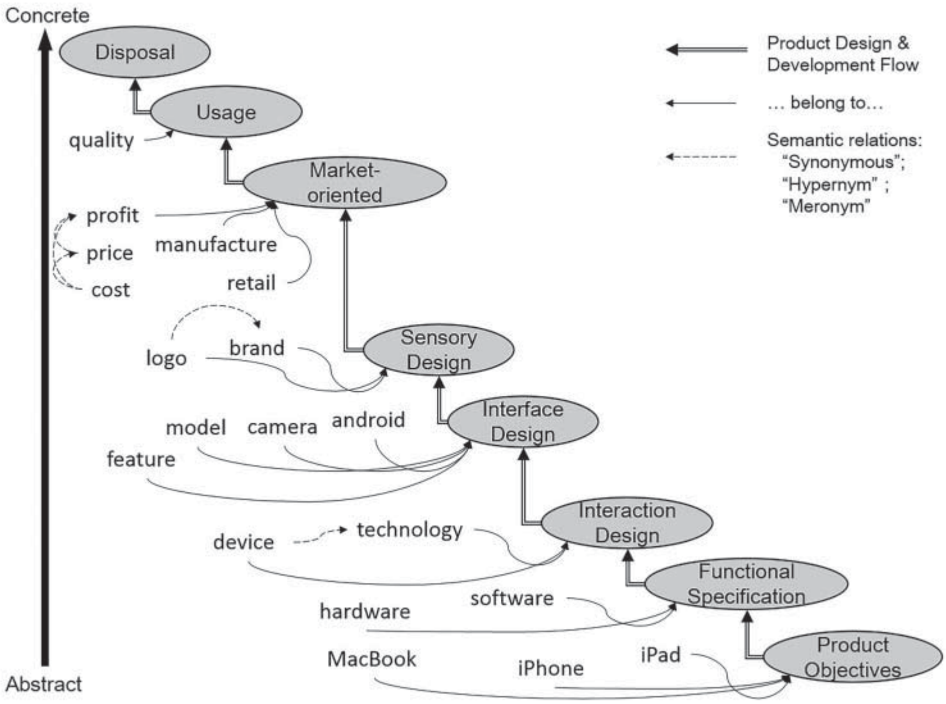

Figure 12. Association of individual tokens of Document 1 referring to PDKH.

help of these two kinds of relations in PDKH, tokens which represent design knowledge can be identified and treated as design tokens.

For sentiment analysis, Text Analysis API, a package of Natural Language Processing, Information Retrieval and Machine Learning tools for extracting meaning and insight from textual and visual content with ease, is introduced to analyze the emotions contained in the collected responses. In general, API can help to analyze the emotions in every sentence and calculate the average sentiment score of all sentences in the document. Therefore, different APIs with different sentiment dictionaries may lead to different results. For this reason, three APIs, i.e. *AYLIEN API for documents*, *AYLIEN API for social media*, and *Meaningcloud API*, are applied to take advantage of multiple sentiment dictionaries in order to achieve relatively accurate estimation of the emotions. A text analytics tool *Meaningcloud* is used to assist in the sentiment analysis, where the external APIs can be imported to train and analyze the sentiment of the target texts. Denoting p_{at_j} as the polarity of token AT_j, the polarities are identified by the three APIs p_{ADat_j}, p_{ASat_j}, and p_{MCat_j}, respectively. The polarity is from the set of {positive, negative, neutral}. The sentiment scores estimated by the three APIs are SS_{ADat_j}, SS_{ASat_j} and SS_{WCat_j}. The sentiment analysis results by these three tools can be processed as follows:

(a) If $p_{ADat_j} = p_{ASat_j}. = p_{MCat_j}$, then $p_{at_j} = p_{ADat_j} = p_{ASat_j} = p_{MCat_j}$, and the sentiment score $SS_{at_j} = \text{Avg}\ (SS_{ADat_j}, SS_{ASat_j}, SS_{WCat_j})$;
(b) If $p_{ADat_j} = p_{ASat_j} \neq p_{MCat_j}$, then $p_{at_j} = p_{ADat_j} = p_{ASat_j}$, and $SS_{at_j} = \text{Avg}\ (SS_{ADat_j}, SS_{ASat_j})$; that is to say, p_{at_j} adopts the polarity of the majority, and SS_{at_j} is the average of the sentiment scores of the majority;

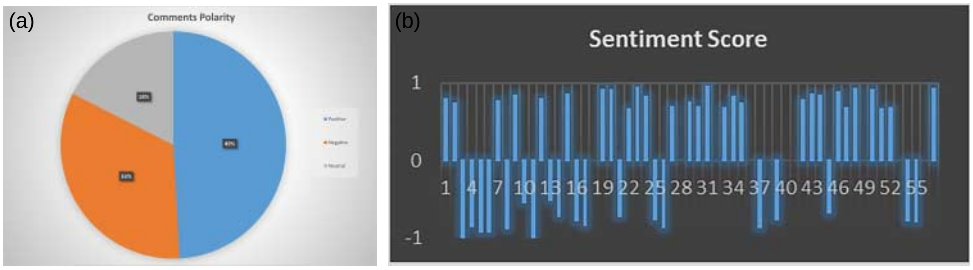

Figure 13. (a) The proportion of documents with different polarities; (b) Sentiment scores of all documents

Table 2. Sentiment analysis results by different APIs.

Sentiment Analysis by APIs		Mean	Standard Error	Number of Documents
Meaningcloud API	Positive	0.86	0.0072	31
	Negative	−0.86	0.0113	13
AYLIEN API for social media	Positive	0.43	0.1312	18
	Negative	−0.79	0.0357	25
AYLIEN API for documents	Positive	0.87	0.0285	32
	Negative	−0.75	0.0527	17
Overall	Positive	0.80	0.0183	28
	Negative	−0.80	0.0304	19

(c) If $p_{ADat_j} \neq p_{ASat_j} \neq p_{MCat_j}$, then expertise is incorporated to identify the p_{at_j}, and SS_{at_j} adopts the value in which the polarity is consistent with expert opinion.

Based on the sentiment analysis by *AYLIEN API for documents*, *AYLIEN API for social media*, and *Meaningcloud API*, the overall sentiment analysis results can be achieved (Table 2). As shown in Figure 13(a,b), although there are more positive responses than negative ones, the difference is not significant. It indicates that consumers stl have dissatisfaction and doubts on this product, which further enhance the need to investigate the emotions in regard to specific design properties.

With the help of APIs, affective tokens can be identified. Two types of affective tokens are examined, i.e. one kind is the affective tokens containing emotions, and another kind involves the concepts/entities which have been identified with emotions in their local context. Some examples of identified affective tokens are listed below (Table 3).

Afterwards, the relationships between design tokens and affective tokens are examined with joint consideration of the semantic relations and lexical reference, and the design tokens which are successfully associated with affective tokens are regarded as product affective properties. The lexical database WordNet is still used to provide connections between different tokens mainly from semantic perspective. As results, more than 400 product affective properties have been identified. Some examples are presented in Table 4.

4.3. Prioritisation of affective design properties

The *DI* and *AI* of each product affective property are calculated using Equations 1–9, and overall priority of product affective properties can be computed using Equations 10 and 11

Table 3. A list of examples of identified affective tokens.

iPhone (N, R1)	Microsoft Corporation (P, R2)	host (P+, R2)	chip (P+, R2)
Apple (N, R1)	LG (P+, R2)	compass (P, R2)	language (N, R3)
Macbook (N, R1)	RAM (P+, R2)	apple (P, R2)	battery (Neu, R3)
Company (P, R1)	Samsung (N, R2)	report (P, R2)	sim (N, R4)
Love (P, R1)	Apple apps (P, R2)	variant (P, R2)	button (N, R4)
High-end (P, R1)	S-AMOLEDS (P+, R2)	telephone (P, R2)	telephone (N, R5)
Innovation (P, R1)	SD (P+, R2)	ecosystem (N, R2)	cover (P, R5)
Product (N, R1)	FHD (N, R2)	competition (P, R2)	accessory (P, R5)
Model (P, R1)	Amazon (P+, R2)	dollar (P, R2)	Internt (N, R6)
iCloud (P, R2)	Google maps (P, R2)	camera (P+, R2)	Wifi (N, R6)
iPhone 5s (P, R2)	update (P, R2)	system (P+, R2)	iPhone (N, R7)
honor (P, R2)	application (P, R2)	standard (P+, R2)	screen (P, R7)
ios (N, R2)	stock (P)	plug (P, R2)	Camera (P, R7)
Android (P+, R2)	feature (P)	interest (P)	Music (P, R7)
Chrome (P, R2)	storage (P)	license (P)	Speaker (N, R8)

Note: The first denotation in brackets represents the polarity (P is positive, N is negative, and NEU is neutral); the second represents the documents from which the token is extracted.

Table 4. Examples of Identified Product Affective Properties (N-grams).

Product Affective Property Examples		
amazed_smoothness	beautiful_designed_products	cheap_headphone
amazed_smoothness_controls	beautiful_device	cheap_water_resistance
amazing_cpu	beautiful_device_look	cheaper_amoled_screen
amazing_cpu_wasted	bigger_battery	chip_powerful
amolded_screen_icing	bigger_battery_play	connector_expandable
amolded_display_pleasing	bigger_battery_speaker	decent_cell_phone
battery_heavy_use	bigger_screen	decent_headphone
battery_issue_timed	buds_excellent	device_simpler
battery_life_excellent	buds_excellent_durable	device_simpler_frustrating
battery_life_great	camera_fantastic	device_audiophile_old
battery_life_lower	camera_fantastic_work	disappointed_fast_charging
battery_life_shorter	camera_faster	dish_pretty_good
battery_life_terrible	camera_faster_processor	display_pleasing
etc ...	etc ...	etc ...

(in this study, *DI* and *AI* are assigned with the same weight, namely, 0.5). To enable a direct comparison and integration of *DI* and *AI*, the calculated results of *DI* and *AI* by Equations 1–9 are normalised into the range of [(0)], (Cross 2000) and denoted as N(*DI*) and N(*AI*). *AI* is the aggregated intensity of all affections, namely, the sum of the absolute values of all positive and negative affections. The final polarity is determined by the algebraic sum of all affections including their polarities. If the algebraic sum is positive, the final polarity is positive, and vice versa. Moreover, the ratio of algebraic sum to *AI* will determine the intensity level of final polarity. If the ratio is significant, it means the final polarity is much stronger than other polarities, and the polarity can be presented by two positive/negative signs.

Considering the large number of product affective properties extracted from online consumer responses, the properties with higher priorities are listed in Table 5 below. Individual design properties are used to represent all the relevant n-gram product affective properties.

According to the results, it was found that the market performance of iPhone 7 is generally good. In particular, the camera, headphone, and battery have the heaviest priorities and significant attention should be paid during the design and improvement process. Especially

Table 5. Product affective properties with higher priority.

Affective Design Properties	Count	DI	N(DI)	AI	N(AI)	OP	Polarity*
Camera	118	86.26	1.000	29.57	0.963	0.982	(++)
Headphone, headphone-jack, headphones, jack	36	26.58	0.255	30.63	1.000	0.627	(+)
Battery	44	32.86	0.333	22.02	0.700	0.517	(−)
Sound, sounds	30	21.74	0.194	19.89	0.626	0.410	(++)
Charging,charge, charged, charger	42	30.98	0.310	14.94	0.453	0.382	(−−)
Look, looks	58	25.97	0.247	13.53	0.404	0.326	(++)
Screen	18	13.33	0.089	11.23	0.324	0.207	(++)
Size	28	19.74	0.169	7.49	0.194	0.181	(++)
Bluetooth	21	14.96	0.109	7.54	0.196	0.153	(−−)
Adapter, adapters	21	15.46	0.116	7.33	0.188	0.152	(−−)
Plug	27	23.31	0.214	4.29	0.082	0.148	(++)
Music	22	19.03	0.160	5.66	0.130	0.145	(++)
Tech, technologies, technology	25	10.96	0.059	4.81	0.101	0.080	(++)
Lightning	10	7.33	0.014	1.95	0.001	0.007	(−)
Software	14	6.20	0.000	1.93	0.000	0.000	(++)

*+ means there are stronger positive responses than negative ones for the property; – means there are stronger negative responses for the property; the number of + and – implies how much this affect is stronger than the other, for example, ++ means there are significantly stronger positive responses about the property.

for the camera, it is the most frequently mentioned property and has received intense emotional feedback from consumers, indicating the necessity to pay attention to camera design so as to ensure users' satisfaction. Actually, camera is always a crucial concern in smart phone design. Almost every generation of iPhone keeps continuously improving its camera. For example, compared with iPhone 6, iPhone 7 has improved the camera resolution from 8MP to 12MP, added optical image stabilisation function and improved flash from 2-tone LED to 4-tone LED. These changes can demonstrate Apple's effort in improving its camera function. However, users' requirements on photographing are varied among different markets. To take Chinese market as an example, selfie is widely loved by users, and photo processing function is greatly desired. For this reason, Huawei has cooperate with Leica. Meitu has launched their own smartphone especially featured by beauty camera. In the respect of camera, iPhone is not the absolutely most competitive one. Nevertheless, it cannot be denied that camera attracts significant consumer concern. In this case study, the consumer data is obtained from Amazon and CNET where English speakers are the main users, and the results show strongly positive emotion on iPhone's camera.

For headphone, it is one of the frequently mentioned product features, but its design importance is not as significant as the affective intensity. It is interesting that the overall emotional polarity of headphone is generally positive (+), rather than strongly positive (++), although it attracts the highest affective intensity. Actually, the comments about headphone are complex, since headphone is related to headphone cable, headphone jack, lightning adapter, wireless headphone, and so on. For users, they normally don't distinguish them distinctly and prefer to use headphone to describe relevant issues. Therefore, the significant changes of iPhone, i.e. the removal of headphone jack, the newly added Lightning-to-3.5-mm adapter, are possibly extracted under this product feature. To reduce such confusion, the case study has narrowed headphone down to the headphone jack and earpods. It was found that consumers give generally positive feedback on 'headphone jack' due to the removal of the 3.5mm headphone jack. However, it is still difficult to avoid all ambiguity. Some negative feedback about lightening adapter and wireless headphone, which are described using headphone related expressions, are also counted.

In general, the positive design properties, such as camera, screen, technology and appearance are the traditional strengths of iPhone and still receive strongly positive comments from consumers in this study. Therefore, the advantages of iPhone 7 on these positive properties should be maintained in future generations.

On the other hand, the battery and charging-related issues are the main pain points of iPhone 7. Referring to the market performance of iPhone 6, the battery problem has been widely complained about and discussed among consumers, so Apple introduced a battery replacement plan for iPhone 6 in a low price (from 79 dollar to 29 dollar). Nonetheless, the battery and charging performance of iPhone 7 is still the focus, and questioned about by consumers. For *adapters*, it can be considered jointly with '*Lightning*' which is also negative, since a new important change for iPhone 7 is the '*Lightning to Headphone Jack Adapter*'. The new change brings some inconvenience to consumers, since they need one more adapter which is easily lost and broken. From the functional perspective, users cannot charge and use headphones simultaneously. That means if the iPhone is in low power and needs charging, the users have to stop all activities using headphones. Moreover, *Bluetooth* is also commented on negatively, which indicates that an improvement on this traditional function may be needed. Actually, the Bluetooth is a classical function of phones. However, the comments about Bluetooth is also complex. The wireless headphone relies on Bluetooth to connect with phone, and the related connection issues are possibly attributed to Bluetooth in users' descriptions. Moreover, with the emergence of wearable products such as Apple watch and other health monitoring devices, they need to be connected to phones via Bluetooth. Therefore, the feedback related to the wireless devices and connection issues are unavoidably mentioned with Bluetooth by users.

To further demonstrate the above results, a comparison between iPhone 7, iPhone 8 and iPhone X is presented. The same configurations are not listed, and the main changes are summarised in Table 6.

Based on the comparison, it can be found that the main improvements of new iPhones generally match the case study results. The camera, battery, charging, Bluetooth and display

Table 6. A summary of main differences between iPhone 6, iPhone 7 and iPhone 8.

	iPhone 7	iPhone 8	iPhone X
Camera	12MP camera	12MP camera	12MP wide-angle and telephoto camerasPortrait mode
Power and Battery	Lasts up to 2 hours longer than iPhone 6s	Lasts about the same as iPhone 7	Lasts up to 2 hours longer than iPhone 7
	–	Wireless charging (Qi-certified chargers)	Wireless charging (Qi-certified chargers)
	–	Fast-charge capable:Up to 50% charge 30 minutes	Fast-charge capable:Up to 50% charge 30 minutes
Cellular and Wireless	Bluetooth 4.2	Bluetooth 5.0	Bluetooth 5.0
Included Headphones	EarPods with Lightning Connector	EarPods with Lightning Connector	EarPods with Lightning Connector
Secure Authentication	Touch ID	Touch ID	Face ID
Display	Retina HD display 4.7-inch (diagonal) widescreen LCD Multi–Touch display with IPS technology	Retina HD display 4.7-inch (diagonal) widescreen LCD Multi–Touch display with IPS technology	Super Retina HD display 5.8-inch (diagonal) all-screen OLED Multi–Touch display

(e.g. screen size) have been importantly improved compared with earlier generations. It is interesting that there is no change to the included headphones of iPhones; however, headphone has been identified as the feature with intensive emotional feedback. It may be caused by that Apple has its headphone product Airpod. From the perspective of marketing and product management strategy, the necessity is not significant to advance the included headphone which is an easily replaceable accessary.

In summary, the usual strengths of iPhone have been further proven, and unsatisfactory properties of iPhone 7 are examined. Since the comments analyzed in this pilot study were posted during the year after iPhone 7 was newly launched, it indicates that not all new changes were well accepted by consumers. Therefore, the analysis on the important affective properties appears particularly important to identify the negative properties and hints at possible improvement directions. Furthermore, the prioritisation provides clear and practical reference for designers as to which property should be paid more attention and how to deploy a further design process.

4.4. Comparison with existing APD studies

In this section, two groups of comparisons are executed: (i) properties extracted from online resources vs. properties evaluated by surveys; and (ii) product affective properties identified based on joint consideration of affect and design vs. based on pure sentiment analysis.

Group 1: *Affective design properties extracted by the proposed approach vs. design properties used in existing APD studies.*

Referring to Table 7, the proposed approach was compared with other APD studies on mobile phones. Through the comparison between the design properties extracted from crowdsourced consumer responses and the design attributes used in existing APD studies (as shown in Table 7), it can be seen that the proposed approach is able to discover more design properties which are of concern by consumers and have provoked consumers' emotions, while existing APD studies are concerned with very limited design attributes. Moreover, these properties cover different product design aspects ranging from hardware, software to marketing issues. It suggests the stronger capability of the proposed approach in unearthing all the potential design concerns significantly influencing consumers' emotion.

Group 2: *Prioritisation of product affective properties based on the joint considerations of design and affective concerns vs. based on only sentiment analysis.*

The prioritisation based on affective intensity and the joint consideration of design and affection are presented in Table 8.

To understand the difference between the two prioritisation results, a focus group was organised. 5 design experts including researchers and engineers with more than 5 years' experience in the industrial design area were involved. They were invited to give evaluation on the design properties according to their expertise and user interaction experience. Generally, the ranking of some design properties is consistent, thus the comparison focuses on the properties with different priorities. In particular, an AHP process is employed. The fundamental AHP scale (1: Equal importance, 3: Moderate importance, 5: Strong importance) is used to compare the properties with different priorities. 1 was assigned to the property with lower priority as basis evaluation score (Table 9).

Table 7. Comparison of the properties extracted by the proposed approach and properties evaluated by surveys.

	General Impression		Hardware							
	Appearance/ Form	Technology	Screen	Sensors	Audio/ Output	Battery	Camera	Storage	SIM Card	Accessory
The proposed approach	√	√	√	√	√	√	√	√	√	√
Yang (2011)	√									
Seva et al. (2011)	√									
Chan et al. (2011)	√									
Fung et al. (2012)	√									
Bhandari, Neben, and Chang (2017)										
Jiang et al. (2015)	√		√							

	Software						Marketing		
	Interface	Phone	Internet Connectivity	Text Input	Email Text Massage	Third-party Application	Payment	Service	delivery
The proposed approach		√	√			√	√	√	√
Yang (2011)									
Seva et al. (2011)									
Chan et al. (2011)									
Fung et al. (2012)									
Bhandari, Neben, and Chang (2017)	√					√			
Jiang et al. (2015)									

Table 8. Prioritisation based on affective concern vs. based on the joint consideration of affective and design concerns.

Prioritisation based on affective intensity (AI)	Prioritisation based on the overall priority (OP)
Headphone, headphones	Camera, cameras
Camera, cameras	Headphone, headphones
Battery, batteries	Battery, batteries
Sound, sounds	Sound, sounds
Screen, screens	Screen, screens
Charging, charge, charged, charger	Charging, charge, charged, charger
Plug	Plug
Music	Tech, technologies,
Tech, technologies,	technology
technology	Music
Adapter, adapters	Adapter, adapters
Look, looks	Bluetooth
Lightning	Look, looks
Bluetooth	Lightning
Size	Size
Software	Software

Table 9. A brief summary of AHP evaluation results.

Camera vs. Headphone	3:1	Camera is one crucial part of iPhone and undoubtedly one of the most cared about functions of users. Therefore, it should be assigned with more priority.
Technology vs. Music	3:1	Technology is one important design aspect. Ofentimes, more effort is devoted to the development of innovative technologies. Moreover, the new technology is always an essential factor to appeal to fans or consumers.
Look vs. Lightning to Headphone Jack Adapter vs. Bluetooth	5:1:3	Industrial design is one success factor of iPhone, thus the appearance/form design is undoubtedly important, no matter for designers or consumers; bluetooth relates to multiple functions such as file transfer and connecting to car bluetooth and other wearable electronics, thus could be assigned with more priorities compared with lightning to headphone jack adapter.

With expertise incorporation, it was found that expert opinions are more consistent with the prioritisation results by the proposed approach. It indicates that the prioritisation based on an integrated consideration of both design and affective concerns is relatively more reasonable. It is promising to provide useful guidance for designers in the design process. Therefore, compared with product property analysis methods from only design perspective, the proposed approach is user/consumer-centered and contains a higher potential to achieve satisfying products. Compared with pure sentiment analysis, the proposed approach involves design concerns and is able to give more comprehensive and rational analysis results. Compared with manual processing of consumers' emotional feedback, the data mining process of the proposed approach is more efficient and scientific.

5. Discussion and conclusions

This work aims to develop a product affective property identification approach. For this end, a web- and text-mining process is deployed to make use of online product review resources, capture useful consumer responses and perform textual analysis. Afterwards, design knowledge hierarchy is constructed to support the identification of design-related tokens, and sentiment dictionaries are utilised to identify affect-related word tokens. With the help of domain ontology and electronic lexical database to provide semantic relations and lexical reference, the associations between design tokens and affective tokens can be examined. The design properties which are related with affective tokens are regarded as affective design properties. The design importance and affective intensity of the affective properties are estimated, and overall priority of these properties can be accordingly achieved. Through a pilot study, it has been demonstrated that the proposed approach is capable of capturing more affective design properties, and a clear and practical reference in terms of the priorities of different design properties can be achieved so as to facilitate decision making and product improvement.

However, there are still some limitations of this work. For example, the identification of affection is based on sentiment analysis, which actually relies on the recognition and measurement of affective words (which have been defined and tagged in existing lexical databases or sentiment dictionaries). If consumers' statements do not include obvious affective words, the feeling may not be detected. Moreover, since the focus of this work is the identification of affective design properties, consumers' emotions are not investigated in very specific types, such as happiness, sadness, confidence and confusion. In future research, consumers' emotions will be further studied. In addition, the redesign strategy based on the product affective features will be explored to facilitate design practice in order to fulfil users' affective expectation and achieve successful product design.

In conclusion, this study explores a product affective properties identification approach based on data mining and is promising in contributing to (i) the discovery of all possible important affective design properties through taking advantage of abundant online consumer responses; (ii) an efficient computation method to deal with large amount of online data, (iii) the integration of design knowledge into affective design so as to achieve more comprehensive and rational understanding of customer emotions, and (iv) prioritisation of affective design properties, which can be practically useful reference to facilitate product design and improvement.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work is supported by Shanghai Pujiang Program [grant number 17PJC054] and also received funding from Shanghai Jiao Tong University.

ORCID

Danni Chang http://orcid.org/0000-0002-6230-5796

References

Adepetu, A., K. A. Ahmed, Y. Al Abd, A. Al Zaabi, and D. Svetinovic. 2012. *CrowdREquire: A Requirements Engineering Crowdsourcing Platform*. Palo Alto, CA: AAAI Spring Symposium Series.

Amiri, M., L. Mohammad-Khanli, and R. Mirandola. 2018. "A Sequential Pattern Mining Model for Application Workload Prediction in Cloud Environment." *Journal of Network and Computer Applications* 105: 21–62.

Bal, A. S., K. Weidner, R. Hanna, and A. J. Mills. 2017. “Crowdsourcing and Brand Control.” *Business Horizons* 60 (2): 219–228.
Balco, P., J. Law, and M. Drahošová. 2017. “Cloud Market Analysis From Customer Perspective.” *Procedia Computer Science* 109: 1022–1027.
Bhandari, U., T. Neben, and K. Chang. 2017. “Wen Yong Chua, Effects of Interface Design Factors on Affective Responses and Quality Evaluations in Mobile Applications.” *Computers in Human Behavior* 72: 525–534.
Bock, C., X. Zha, H. W. Suh, and J. H. Lee. 2010. “Ontological Product Modeling for Collaborative Design.” *Advanced Engineering Informatics* 24 (4): 510–524.
Boudreau, K. J., and K. Lakhani. 2013. “Using the Crowd as an Innovation Partner.” *Harvard Business Review* 91 (4): 60–69.
Cai, J., J. Luo, S. Wang, and S. Yang. 2018. “Feature Selection in Machine Learning: A New Perspective.” *Neurocomputing*. ISSN 0925-2312, doi:10.1016/j.neucom.2017.11.077.
Cao, D., Z. Li, and K. Ramani. 2011. “Ontology-Based Customer Preference Modeling for Concept Generation.” *Advanced Engineering Informatics* 25 (2): 162–176.
Chan, K. Y., C. K. Kwong, T. S. Dillon, and K. Y. Fung. 2011. “An Intelligent Fuzzy Regression Approach for Affective Product Design That Captures Nonlinearity and Fuzziness.” *Journal of Engineering Design* 22 (8): 523–542.
Chang, D., C.-H. Chen, and K. M. Lee. 2014. “A Crowdsourcing Development Approach Based on a Neuro-Fuzzy Network for Creating Innovative Product Concepts.” *Neurocomputing* 142: 60–72.
Chen, X. Y., C. H. Chen, K. F. Leong, and X. Jiang. 2013. “An Ontology Learning System for Customer Needs Representation in Product Development.” *The International Journal of Advanced Manufacturing Technology* 67: 441–453.
Chen, L.-C., and P.-Y. Chu. 2012. “Developing the Index for Product Design Communication and Evaluation From Emotional Perspectives.” *Expert Systems with Applications* 39: 2011–2020.
Chen, J. L., and C.-W. Liau. 2001. “A Simple Life Cycle Assessment Method for Green Product Conceptual Design.” In *Proceedings Second International Symposium on Environmentally Conscious Design and Inverse Manufacturing*, 11–15 December, Tokyo, Japan, 775–780.
Chuang, M. C., C. C. Chang, and S. H. Hsu. 2001. “Perceptual Factors Underlying User Preferences Toward Product Form of Mobile Phones.” *International Journal of Industrial Ergonomics* 27: 247–258.
Crilly, N., J. Moultrie, and J. Clarkson. 2004. “Seeing Things: Consumer Response to the Visual Domain in Product Design.” *Design Studies* 25 (6): 547–577.
Cross, N. 2000. *Engineering Design Methods: Strategies for Product Design*. 3rd ed.Chichester, UK: Wiley.
de Mattos, C. A., K. O. Kissimoto, and F. J. B. Laurindo. 2018. “The Role of Information Technology for Building Virtual Environments to Integrate Crowdsourcing Mechanisms Into the Open Innovation Process.” *Technological Forecasting & Social Change*. doi:10.1016/j.techfore.2017.12.020.
Diego-Mas, J. A., and J. Alcaide-Marzal. 2016. “Single Users’ Affective Responses Models for Product Form Design.” *International Journal of Industrial Ergonomics* 53: 102–114.
Djelassi, S., and I. Decoopman. 2013. “Customers’ Participation in Product Development Through Crowdsourcing: Issues and Implications.” *Industrial Marketing Management* 42 (5): 683–692.
Djenouri, Y., A. Belhadi, and R. Belkebir. 2018. “Bees Swarm Optimization Guided by Data Mining Techniques for Document Information Retrieval.” *Expert Systems With Applications* 94: 126–136.
Ehrig, M., P. Haase, M. Hefke, and N. Stojanovic. 2005. “Similarity for Ontologies – A Comprehensive Framework.” In *Proceedings of 13th European Conference on Information Systems (ECIS)*, May 26–28, Regensburg, Germany, 127–138.
Fung, K. Y., C. K. Kwong, K. W. M. Siu, and K. M. Yu. 2012. “A Multi-Objective Genetic Algorithm Approach to Rule Mining for Affective Product Design.” *Expert Systems with Applications* 39: 7411–7419.
Garrett, J. J. 2011. *The Elements of User Experience: User-Centered Design for the Web and Beyond*. 2nd ed. Berkeley: Pearson Education.
Gennaria, R., A. Melonioa, D. Raccanellob, M. Brondino, G. Dodero, M. Pasinib, and S. Torello. 2017. “Children’s Emotions and Quality of Products in Participatory Game Design.” *International Journal of Human-Computer Studies* 101: 45–61.
Hayes, R. H., S. C. Wheelwright, and K. B. Clark. 1988. *Dynamic Manufacturing*. New York: Free Press.

He, W. 2013. "Improving User Experience with Case-Based Reasoning Systems Using Text Mining and Web 2.0." *Expert Systems with Applications* 40 (2): 500–507.

Hedges, M., and S. Dunn. 2018. "Chapter 3 - Processes and Products: A Typology of Crowdsourcing." In *Academic Crowdsourcing in the Humanities*, 27–49. Chandos Publishing. ISBN 9780081009413. doi:10.1016/B978-0-08-100941-3.00003-6.

Howe, J. 2006. "The Rise of Crowdsourcing." *WiredMag* 14 (6): 1–5.

Hsiao, S. W., and H. C. Tsai. 2005. "Applying a Hybrid Approach Based on Fuzzy Neural Network and Genetic Algorithm to Product Form Design." *International Journal of Industrial Ergonomics* 35: 411–428.

Huang, Y., C.-H. Chen, and L. P. Khoo. 2012. "Kansei Clustering for Emotional Design Using a Combined Design Structure Matrix." *International Journal of Industrial Ergonomics* 42: 416–427.

Huang, Y., C.-H. Chen, and L. P. Khoo. 2012. "Products Classification in Emotional Design Using a Basic-Emotion Based Semantic Differential Method." *International Journal of Industrial Ergonomics* 42: 569–580.

Huang, Y., A. Shema, and H. Xia. 2017. "A Proposed Genome of Mobile and Situated Crowdsourcing and its Design Implications for Encouraging Contributions." *International Journal of Human-Computer Studies* 102: 69–80.

Hwang, K., X. Bai, M. Shi, Y. Li, W. G. Chen, and Y. Wu. 2016. "Cloud Performance Modeling and Benchmark Evaluation of Elastic Scaling Strategies." *IEEE Transactions on Parallel and Distributed Systems* 27 (1): 130–143.

Jiang, H., C. K. Kwong, K. W. M. Siu, and Y. Liu. 2015. "Rough set and PSO-Based ANFIS Approaches to Modeling Customer Satisfaction for Affective Product Design." *Advanced Engineering Informatics* 29: 727–738.

Jordan, P. W. 1998. "Human Factors for Pleasure in Product use." *Applied Ergonomics* 29: 25–33.

Khalid, H. 2006. "Embracing Diversity in User Needs for Affective Design." *Applied Ergonomics* 37 (4): 409–418.

Kwong, C. K., H. Jiang, and X. G. Luo. 2016. "AI-based Methodology of Integrating Affective Design, Engineering, and Marketing for Defining Design Specifications of New Products." *Engineering Applications of Artificial Intelligence* 47: 49–6.

Lai, H., Y. Chang, and H. Chang. 2005. "A Robust Design Approach for Enhancing the Feeling Quality of a Product: A Car Profile Case Study." *International Journal of Industrial Ergonomics* 35: 445–460.

Lane, N. D., E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell. 2010. "A Survey of Mobile Phone Sensing." *IEEE Communications Magazine* 48 (9): 140–150.

Lau, R. Y. K., D. Song, Y. Li, C. H. Cheung, and J. X. Hao. 2009. "Towards a Fuzzy Domain Ontology Extraction Method for Adaptive E-Learning." *IEEE Transactions on Knowledge and Data Engineering* 21 (6): 800–813.

Lee, C. Y., B. Song, and Y. T. Park. 2012. "Design of Convergent Product Concepts Based on Functionality: An Association Rule Mining and Decision Tree Approach." *Expert Systems with Applications* 39 (10): 9534–9542.

Li, S.-T., and F.-C. Tsai. 2013. "A Fuzzy Conceptualization Model for Text Mining with Application in Opinion Polarity Classification." *Knowledge-Based Systems* 39: 23–33.

Lokman, A. M., and K. A. Kamaruddin. 2010. Kansei Affinity Cluster for Affective Product Design, 2010 International Conference on User Science Engineering (i-USEr).

López-Sánchez, D., A. G. Arrieta, and J. M. Corchado. 2018. "Deep Neural Networks and Transfer Learning Applied to Multimedia Web Mining." In *Distributed Computing and Artificial Intelligence, 14th International Conference. DCAI 2017. Advances in Intelligent Systems and Computing, Vol 620*, edited by S. Omatu, S. Rodríguez, G. Villarrubia, P. Faria, P. Sitek, and J. Prieto, 124–131. Cham: Springer.

Lu, W., and J.-F. Petiot. 2014. "Affective Design of Products Using an Audio-Based Protocol: Application to Eyeglass Frame." *International Journal of Industrial Ergonomics* 44: 383–394.

Lv, Y. Q., K. M. Lee, and Z. Wu. 2011. "Fuzzy Theory Applied in Quality Management of Distributed Manufacturing System: A Literature Review and Classification." *Engineering Applications of Artificial Intelligence* 24 (2): 66–277.

Ma, J., W. Xu, E. Turban, S. Wang, and O. Liu. 2012. "An Ontology Based Text Mining Method to Cluster Proposals for Research Project Selection." *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 42 (3): 784–790.

Maedche, A., and S. Staab. 2001. "Ontology Learning for the Semantic Web." *IEEE Intelligent Systems* 16 (2): 72–79.

Mazhar, M. I., S. Kara, and H. Kaebernick. 2007. "Remaining Life Estimation of Used Components in Customer Products: Life Cycle Data Analysis by Weibull and Artificial Neural Networks." *Journal of Operations Management* 25 (6): 1184–1193.

Nagamachi, M. 2002. "A Fuzzy Rule-Based Approach to Modeling Consumer-Oriented Technology for Product Development." *Applied Ergonomics* 33: 289–294.

Nazari-Shirkouhi, S., and A. Keramati. 2017. "Modeling Customer Satisfaction with New Product Design Using a Flexible Fuzzy Regression-Data Envelopment Analysis Algorithm." *Applied Mathematical Modelling* 50: 755–771.

Negnevitsky, M. 2011. *Artificial Intelligence: A Guide to Intelligent Systems*. 3rd ed. London: Pearson Education Limited.

Newman, R., V. Chang, R. J. Walters, and G. B. Wills. 2016. "Web 2.0—The Past and the Future." *International Journal of Information Management* 36 (4): 591–598.

Osgood, C., G. Suci, and P. Tannenbaum. 1967. *The Measurement of Meaning*. Urbana-Champaign, IL: University of Illinois Press.

Park, J., and S. H. Han. 2004. "A Fuzzy Rule-Based Approach to Modeling Affective User Satisfaction Towards Office Chair Design." *International Journal of Industrial Ergonomics* 34: 31–47.

Park, J.-H., and K.-K. Seo. 2003. "A Knowledge-Based Approximate Life Cycle Assessment System for Evaluating Environmental Impacts of Product Design Alternatives in a Collaborative Design Environment." *Advanced Engineering Informatics* 20 (2): 147–154.

Pengnate, S., and R. Sarathy. 2017. "An Experimental Investigation of the Influence of Website Emotional Design Features on Trust in Unfamiliar Online Vendors." *Computers in Human Behavior* 67: 49–60.

Poria, S., and E. Cambria. 2016. "Alexander Gelbukh, Aspect Extraction for Opinion Mining with a Deep Convolutional Neural Network." *Knowledge-Based Systems* 108: 42–49.

Prado-León, L. R. 2015. "Color Preferences in Household Appliances: Data for Emotional Design." *Procedia Manufacturing* 3: 5707–5714.

Rajpathak, D. G. 2013. "An Ontology Based Text Mining System for Knowledge Discovery From the Diagnosis Data in the Automotive Domain." *Computers in Industry* 64 (5): 565–580.

Serra, I., R. Girardi, and P. Novais. 2014. "Evaluating Techniques for Learning non-Taxonomic Relationships of Ontologies From Text." *Expert Systems with Applications* 41 (11): 5201–5211.

Seva, R. R., H. B. L. Duh, and M. G. Helander. 2007. "The Marketing Implications of Affective Product Design." *Applied Ergonomics* 38: 723–731.

Seva, R. R., K. G. T. Gosiaco, M. C. Santos, and D. M. Pangilinan. 2011. "Product Design Enhancement Using Apparent Usability and Affective Quality." *Applied Ergonomics* 42: 511–517.

Shi, L., X. Ma, L. Xi, Q. Duan, and J. Zhao. 2011. "Rough set and Ensemble Learning Based Semi-Supervised Algorithm for Text Classification." *Expert Systems with Applications* 38: 6300–6306.

Song, W., C. H. Li, and S. C. Park. 2009. "Genetic Algorithm for Text Clustering Using Ontology and Evaluating the Validity of Various Semantic Similarity Measures." *Expert Systems with Applications* 36 (5): 9095–9104.

Sousa, I., and D. Wallace. 2006. "Product Classification to Support Approximate Life-Cycle Assessment of Design Concepts." *Technological Forecasting and Social Change* 73 (3): 228–249.

Stumme, G., A. Hotho, and B. Berendt. 2006. "Semantic Web Mining State of the art and Future Directions." *World Wide Web* 4: 124–143.

Tsai, C.-W., S.-J. Liu, and Y.-C. Wang. 2017. "A Parallel Metaheuristic Data Clustering Framework for Cloud." *Journal of and Distributed Computing*. doi:10.1016/j.jpdc.2017.10.020.

Wang, X., M. Hauswirth, T. Vitvar, and M. Zaremba. 2008. "Semantic Web Services Selection Improved by Application Ontology with Multiple Concept Relations." In *Proceedings of the 23rd Annual ACM Symposium on Applied Computing*, 16–20 March, Ceara, Brazil, 2237–2242.

Wang, J., J. Peng, and O. Liu. 2015. “A Classification Approach for Less Popular Webpages Based on Latent Semantic Analysis and Rough set Model.” *Expert Systems with Applications* 42: 642–648.

Wang, S., Z. Zhang, K. Ye, H. Wang, and X. Chen. 2008. “An Ontology for Causal Relationships Between News and Financial Instruments.” *Expert Systems with Applications* 35: 569–580.

Yan, W., C.-H. Chen, Y. Huang, and W. Mi. 2009. “A Data-Mining Approach for Product Conceptualization in a Web-Based Architecture.” *Computers in Industry* 60 (1): 21–34.

Yang, C.-C. 2011. “Constructing a Hybrid Kansei Engineering System Based on Multiple Affective Responses: Application to Product Form Design.” *Computers & Industrial Engineering* 60: 760–768.

Yang, C.-C., and M.-D. Shieh. 2010. “A Support Vector Regression Based Prediction Model of Affective Responses for Product Form Design.” *Computers & Industrial Engineering* 59: 682–689.

Yeh, I.-C., C.-H. Lien, T.-M. Ting, and C.-H. Liu. 2009. “Applications of Web Mining for Marketing of Online Bookstores.” *Expert Systems with Applications* 36: 11249–11256.

Yung, D., M. L. Li, and S. K. Chang. 2014. “Evolutionary Approach for Crowdsourcing Quality Control.” *Journal of Visual Languages and Computing* 25: 879–890.

Zhai, L.-Y., L.-P. Khoo, and Z.-W. Zhong. 2009. “A Rough Set Based Decision Support Approach to Improving Consumer Affective Satisfaction in Product Design.” *International Journal of Industrial Ergonomics* 39: 295–302.

Zhang, Q., L. T. Yang, Z. Chen, P. Li, and M. J. Deen. 2017. “Privacy-Preserving Double-Projection Deep Computation Model with Crowdsourcing on Cloud for Big Data Feature Learning.” *IEEE Internet of Things Journal* 99: 1–1.

Zolkifly, N. H., and S. N. Baharom. 2016. “Selling Cars Through Visual Merchandising: Proposing Emotional Design Approach.” *Procedia Economics and Finance* 37: 412–417.