IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 1, 2018, accepted February 6, 2018, date of publication February 22, 2018, date of current version April 18, 2018.

*Digital Object Identifier 10.1109/ACCESS.2018.2807622*

# Innovation Topic Analysis of Technology: The Case of Augmented Reality Patents

**HAYOUNG CHOI[1], SEUNGHYUN OH[1], SUNGCHUL CHOI[2], AND JANGHYEOK YOON[1]**
[1]Department of Industrial Engineering,KonkukUniversity, Seoul 05029, South Korea
[2]Department of Industrial and Management Engineering, GachonUniversity, Seongnam 13120, South Korea

Corresponding author: Janghyeok Yoon (janghyoon@konkuk.ac.kr)

This paper was supported by Konkuk University in 2015.

**ABSTRACT** Innovation topics within a technology can be defined as generalized technical subjects that the technology desires to develop or improve. Since innovation topics act as a driving force for technical innovation, monitoring such concepts is necessary for understanding the current technology and directing further R&D. However, little attention has been paid to identifying latent innovation topics within a technology, and analyzing their relationships and potential for opportunity. Therefore, this paper proposes a multi-step approach to technological innovation topic analysis, on the basis of patents. The steps consist of: 1) structuring patent-keyword vectors; 2) identifying innovation topics based on semantic patent analysis; 3) constructing an innovation topic network; and 4) generating an opportunity-focused innovation topic map. The process of our approach is illustrated using patents that are related to augmented reality. This method can contribute to the systematic monitoring of a technology system's innovation topics and their potential.

**INDEX TERMS** Patent mining, topic modeling; network analysis, opportunity analysis, augmented reality technology.

## I. INTRODUCTION

Innovation is regarded as the application of better solutions that satisfy new requirements, unarticulated needs, or existing markets [1], and thus it is considered the lifeblood of technology-based firms for their sustainable and successful business [2]. For this reason, such firms attempt to manage technology innovation activities through strategic R&D processes [3], in which determining the development directions for a particular target technology precedes all other tasks [4]. Therefore, an effective understanding of existing and new/emerging technology innovation topics in a technology system can assist experts in directing further R&D.

Innovation concepts refer to abstract or generalized ideas that encourage and propagate innovation [5]. From a technical point of view, innovation concepts are categorized into generalized technical topics or subjects that a given technology attempts to develop or improve. For example, a battery technology may have existing technical innovation topics, such as durability, recharging, and lifetime, as well as new or emerging innovation topics, such as solar-powered and wireless charging. Therefore, technical innovation topics within a technology are the driving force of innovation activities of technology-based firms and have the ability to describe the focal features targeted by the past and current technology. For this reasons, identifying innovation topics and their characteristics, such as importance and potential, in a given technology can assist technology experts in understanding the past and current of a technology and determining its further directions for R&D.

Patents, each of which contains the entirety of information on its relevant invention, serve as effective materials for identifying technical innovation topics. In particular, patent analysis can appropriately describe the evolution process of technologies in various industries [6], because developments in both academia and industry require patent applications to secure their R&D results under the patent system [7]. In addition, every patent, regardless of its commercial value, is a final output of R&D and can thus be used to generate technical insight into subsequent technology development [8]. For these reasons, patents have been widely used as the primary material for identifying trends in rapidly evolving technologies. Invention contents are mainly described in the textual sections of patents, such as abstracts, detailed descriptions, and claims. Therefore, a number of patent-based studies for technology monitoring have been conducted to structure and visualize the technical data obtained from patents, and analyze the structured data indices. Methods that have been used are patent-level mapping, including patent maps [9], [10]; patent networks [11]–[13]; dynamic patent lattices [14]; term or syntactic structure-level mapping, including keyword networks [15]; invention property-function networks [16], [17]; and key graphs [18].

Despite the contributions of these studies, they have exhibited limitations in terms of technology innovation topic analysis. The first limitation is a result of patent or term-level analysis. Prior studies have used clustering techniques based on patent or term relatedness to identify technology clusters, but could not successfully identify and analyze the latent technical subjects underlying multiple patents or terms. For example, in these methods, it is assumed that a patent belongs to only one technology subject, despite the possibility of its belonging to multiple subjects. In addition, most studies have analyzed primary or emerging technology clusters, without considering their opportunity potential. Examining technology subject potential based on their importance and satisfaction could provide insightful information for R&D decision making. Therefore, consideration of both technical innovation topics and their opportunity has remained unexplored thus far.

Therefore, we propose a multi-step approach to monitoring technical innovation topics from patents, using latent Dirichlet allocation (LDA)-based topic modeling and social network analysis (SNA). LDA is a technique for topic modeling in which a process is applied to discover latent or abstract topics occurring in a collection of documents, while SNA is a technique for mapping and measuring the relationships among connected entities. Finally, in order to determine future R&D opportunities, innovation topics are evaluated according to the concept of opportunity, based on importance and satisfaction. This approach includes: 1) structuring patents based on a vector space model (VSM), 2) identifying innovation topics and their relationships based on LDA, and 3) constructing innovation topic networks and generating opportunity-focused innovation topic maps. The workings of this approach are illustrated using patents relating to augmented reality technology, which is an emerging technology that has recently achieved a rapid growth through its various industrial applications and that requires further development directions for R&D.

The contributions of this study are threefold. Firstly, this method enables innovation topic-focused technology analysis. While previous approaches addressed information mapping and its interpretation based on the similarity and relatedness of patents or terms, this study provides a different view of technology analysis, by focusing on technical innovation topics that the system in question is attempting to develop or improve. Secondly, this study can support technology experts in determining future R&D directions, because innovation topics and their potential are analyzed from an opportunity perspective. Finally, our quantified approach can be an effective aid for monitoring innovation topics and their changing panorama in the rapidly evolving high and emerging technologies. This is because this method can be performed computationally, and is neutral in terms of the type of technology analyzed.

This paper continues by describing an overview of the groundwork. Then, we present the proposed innovation topics analysis approach and its application for technology planning. Finally, we conclude the paper with a discussion and an outline of future research topics.

## II. THEORETICAL BACKGROUND

Our approach is based on VSM, LDA, and SNA; therefore, this section provides a brief overview of these theoretical backgrounds.

### A. VECTOR SPACE MODEL(VSM) IN PATENT ANALYSIS

The VSM is an algebraic model for representing a document as an array of identifiers, such as, index terms, for example [19]. Since this model can disambiguate documents as well as any entity, it has been widely used in information retrieval [20] and text clustering [21]. In this method, each text document is expressed as a vector of terms, and a vectorized document is composed of term identifiers, each of whose value may be a weighting according to its significance. One of the most effective means of computing a term's weighting is term frequency-inverse document frequency (TF-IDF). This is a weighing factor as well as numerical statistic that reflects how important a term is to a document in a corpus [22]. The underlying concept of TF-IDF is that a document's terms can be divided into those with eliteness and those without [23]. Therefore, TF-IDF has a value proportional to the number of times a word appears in a document, and also counterbalances the word's frequency in a corpus (Eq. 1).

$$w_{i,j} = tf_{i,j} \times \log\left(\frac{N}{df_i}\right), \tag{1}$$

where $w_{i,j}$ is term $i$'s weighting in document $j$, $tf_{i,j}$ is the number of occurrences of term $i$ in document $j$, $df_i$ is the number of documents containing term $i$, and $N$ is the total number of documents in the corpus.

Due to the applicability of structuring text documents, patent analysis studies have employed the VSM to measure relatedness among patents or keywords and generate patent maps or networks. In this way, patent vacuums, core patents, and technology clusters have been analyzed to assist experts in technology planning processes, such as technology trend analysis [13], [24] and new technology opportunity identification [25]. As the VSM has been widely used for structuring documents, in order to extract innovation topics from patent texts based on topic modeling, this study makes use of the VSM to represent each patent as a vector of keywords found in that patent, along with their weightings.

### B. LATENT DIRICHLET ALLOCATION(LDA)

LDA is a generative model that determines topics from a set of documents based on term frequency [26]. More precisely, LDA automates the topic discovery process, considering documents as a mixture of topics that categorize words with certain probabilities. Therefore, LDA produces two final outputs: individual document topic distributions and topic word distributions.

The underlying assertion of LDA is that documents are presented as a random mixture covering latent topics, wherein each topic is characterized by a distribution over words [27]. LDA assumes the following generative process for a corpus $D$ consisting of $K$ topics and $M$ documents, each of length $N_i$:

1. Choose $\theta_i \sim Dir(\alpha)$, where $i \in \{1, \ldots, M\}$;
2. Choose $\varphi_k \sim Dir(\beta)$, where $k \in \{1, \ldots, K\}$;
3. For each word position i, j, where $j \in \{1, \ldots, N_i\}$ and $i \in \{1, \ldots, M\}$:
   - Choose a topic $z_{ij} \sim Multinomial(\theta_i)$
   - Choose a word $w_{ij} \sim Multinomial(\varphi_{z_{ij}})$.

In the above process, $\alpha$ is the Dirichlet prior parameter in the per-document topic distributions, $\beta$ is the Dirichlet prior parameter in the per-topic word distribution, $\theta_i$ is the topic distribution for document $i$ (the sum of $\theta_i$ is 1.0), $\varphi_k$ is the word distribution for topic $k$, $z_{ij}$ is the topic of the $j^{th}$ word in document $i$, and $w_{ij}$ is the specific word.

Due to the advantages offered by LDA for textual analysis, a number of studies have used this method for the applications of web spam filtering [28], fraud detection [29], and scientific article and web site recommendation [27], [30]–[32]. In terms of patent analysis, certain studies have used LDA-based topic modeling for generating patent development maps [7], mapping technological knowledge landscapes [33], and identifying product opportunities [34]. In this study, we use LDA to identify innovation topics underlying massive patents and their relationships, in order to generate input for innovation topic networks.

### C. SOCIAL NETWORK ANALYSIS(SNA)

A social network consists of nodes (actors) and links (relationships) connecting the nodes. SNA enables the visualization and measurement of the relationships and interactions among any objects, such as people, groups, organizations, computers, and other connected entities [35]. The visual produced by such network analysis provides an overall understanding of the nodes and their relationships within a network; however, as the number of nodes increases, it becomes more difficult to understand the network. Therefore, network indicators are beneficial for understanding the specific characteristics of nodes and clusters within a network [36]. For example, a node's centrality indicates its relative importance within a network, while the density of a network or its subnetwork illustrates how closely all nodes in the network are related.

In its early stages, SNA emerged as a key modern sociology technique to study relationship patterns among social actors; however, its application areas have also expanded into technology analysis based on technical data, including patents and journal papers. In particular, patent-based approaches have constructed technology networks using patent citations and classifications, to measure inter-industrial knowledge flows [37], [38], identify core and emerging fields in rapidly evolving technologies [12], [39], and map technological trajectories [40]. In this study, we construct a network based on the various innovation topic co-occurrences extracted from a given technology's patents, thereby visualizing the relationship among topics and measuring their relative importance or impact in the technology system.

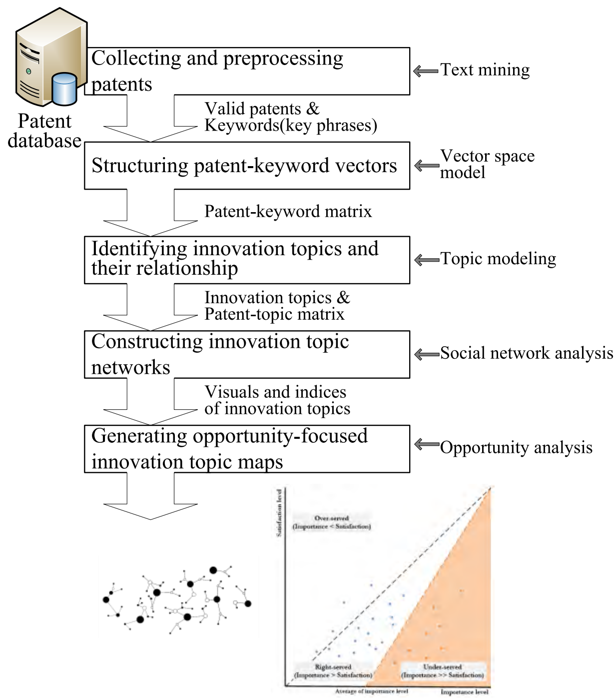



**FIGURE 1. Overall procedure.**

## III. PROPOSED APPROACH

Our innovation topic analysis approach consists of five steps (Figure 1): 1) collecting and preprocessing patents related to a given technology; 2) structuring patents as keyword (or key phrase) vectors using VSM, 3) extracting semantic innovation topics and their relationships by means of LDA-based topic modeling, 4) constructing innovation topic networks; and 5) generating an opportunity-focused innovation topic (OFIT) map. These steps are described in greater detail in the following subsections.

### A. COLLECTING AND PREPROCESSING OF PATENTS

The first step of our approach involves collecting and preprocessing a set of patents relating to a given technology. The collection process consists of defining taxonomies for the technology under study and constructing a patent retrieval query statement with the keywords corresponding to these taxonomies. Then, the patents for analysis can be located using online patent database services, such as KIPRIS (http://www.kipris.or.kr) and WIPSON (http://www.wipson.com). The patents returned by the retrieval query can be stored in an electronic file format, including Microsoft Excel and text files.

Once the patents for analysis have been determined, keywords are extracted from each of them. Generally, a patent's text is a document consisting of multiple sentences, each of which in turn contains keywords, including single or compound words. Such keywords can be extracted from the text by means of natural language processing (NLP) tools, such as RAKE (https://github.com/aneesha/RAKE), AlchemyAPI (http://www.alchemyapi.com), Aylien (http://aylien.com/text-api), and TextRazor (https://www.textrazor.com/docs/rest). However, certain extracted keywords, such as "system" and "process," should be excluded from the list, because they may be irrelevant or too general for textual analysis. Finally, by excluding such irrelevant keywords, a set of valid keywords is prepared for structuring patents in the next step.

### B. STRUCTURING PATENTS AS KEYWORD VECTORS

The aim of this step is to generate a patent-keyword matrix. To this end, each patent is first structured as a patent-keyword vector, or an array of keywords (or key phrases) appearing in the patent, based on VSM. Each element of a patent-keyword vector represents the occurrence frequencies of its corresponding keyword in the patent, and we can obtain an initial matrix by incorporating the patent-keyword vectors of all patents. However, using only occurrence frequencies results in a bias towards general or common technical keywords, which means that rare but important keywords may be undervalued. For example, if some certain keywords with a high frequency value, such as "HMD," "signal," "capture device," and "gesture," are highly common across all patents, they may not provide sufficient quality to distinguish between relevant and non-relevant patents.

For this reason, during this step, the weighting value of each matrix element is identified based on the concept of TF-IDF (Eq. 1). In terms of TF-IDF, a technical keyword reaches a high weight by having a high term frequency in its corresponding patent and a low document frequency in the entire patent collection. Therefore, a weighted patent-keyword matrix based on the TF-IDF concept can describe the uniqueness, generality, and importance of each technical keyword in the technology under consideration (Figure 2). The output matrix obtained is used to determine the innovation topics from patents and the relationships among these in the following steps.

### C. IDENTIFYING INNOVATION TOPICS AND THEIR RELATIONSHIPS

In this step, LDA-based topic modeling is adopted to identify the innovation topics portrayed in multiple patents. Our topic modeling application requires two inputs: the patent-keyword matrix and number of topics. The patent-keyword matrix obtained from the previous step can be used as the input matrix, while an appropriate number of topics should be determined for topic modeling. Among several techniques available for selecting the number of topics, we opted to use the elbow method. With this method, an optimal number of topics is determined by the average cosine similarity between all pairs of topic-keyword distribution vectors outputted by the topic modeling; the number of topics with the lowest cosine similarity is selected as the optimal number that distinguishes the semantic topics to the greatest degree [41].

Our topic modeling application produces two output matrices: a topic-keyword distribution and patent-topic distribution matrix (Figure 2). Firstly, each row vector of the topic-keyword distribution matrix describes how a topic is constructed by its main contributing keywords, so each topic's labeling is conducted based on these keywords and their probability of contribution to the topic. For example, if a topic has main keywords such as "sound," "sound volume," "headphone," and "speaker," this topic can be named "sound" or "product sound." The labeled topics in this study are considered as innovation topics, because they are the technical subjects that are currently targeted by patents in the system under consideration.

Secondly, each row vector of the patent-topic distribution matrix indicates how the identified innovation topics contribute to the construction of each patent; thus, it can be used to determine to which what innovation topics a patent belongs. However, because each patent has a portion of innovation topics in the form of probability, a relative threshold value $\theta$ is used for each patent, so that the belongingness of the patent to the topics is represented by a binary value of 0 or 1. To this end, we modify the empirical method for identifying the optimized cut-off value [42], in order to select an appropriate threshold for each patent's topic determination. The method for cut-off value optimization is based on the VSM similarity calculation between a standard adjacency VSM, in the form of a vector or matrix, and its pseudo adjacency VSMs (Figure 3). The adjacency VSM is created by its entry values; that is, entries are converted into 1 if the entry value is not 0, and into 0 if it is 0. The standard adjacency VSM is accordingly created in terms of each patent-topic distribution vector, and the pseudo adjacency VSMs are generated in terms of the normalized patent-topic vector, by selecting normalized cut-off values from 0 to 1, with an accumulation interval of 0.01. The similarity between the standard adjacency VSM and its pseudo VSMs is calculated by means of a simple match coefficient, as follows

$$S_{i,j} = (a + d) / (a + b + c + d), \tag{2}$$

where $a$ is the number of entries that are 1 in both the standard and pseudo adjacency VSM, $b$ is the number of entries that are 1 in the standard and 0 in the pseudo adjacency VSM, $c$ is the number of entries that are 0 in the standard and 1 in the pseudo adjacency VSM, and $d$ is the number of entries that are 0 in both the standard and pseudo adjacency VSM. Finally, 95% of the maximum match between the pseudo adjacency VSMs and a standard adjacency VSM is selected as the optimal cut-off value [42].

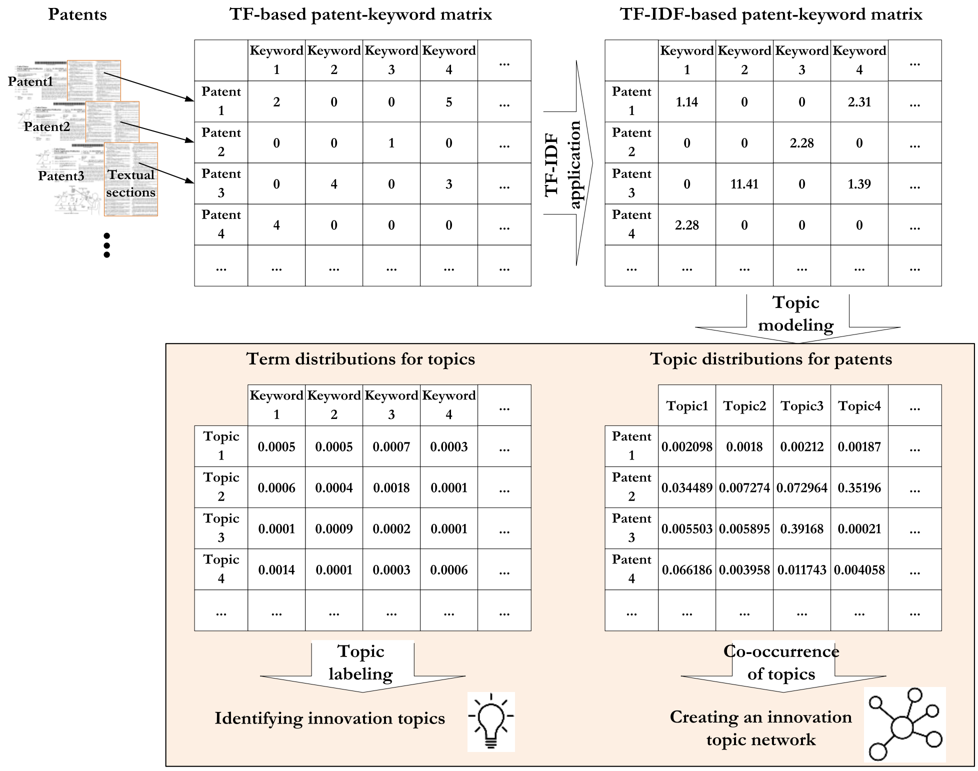



**FIGURE 2.** Concept of patent-keyword matrix generation and output matrices of LDA-based topic modelling.

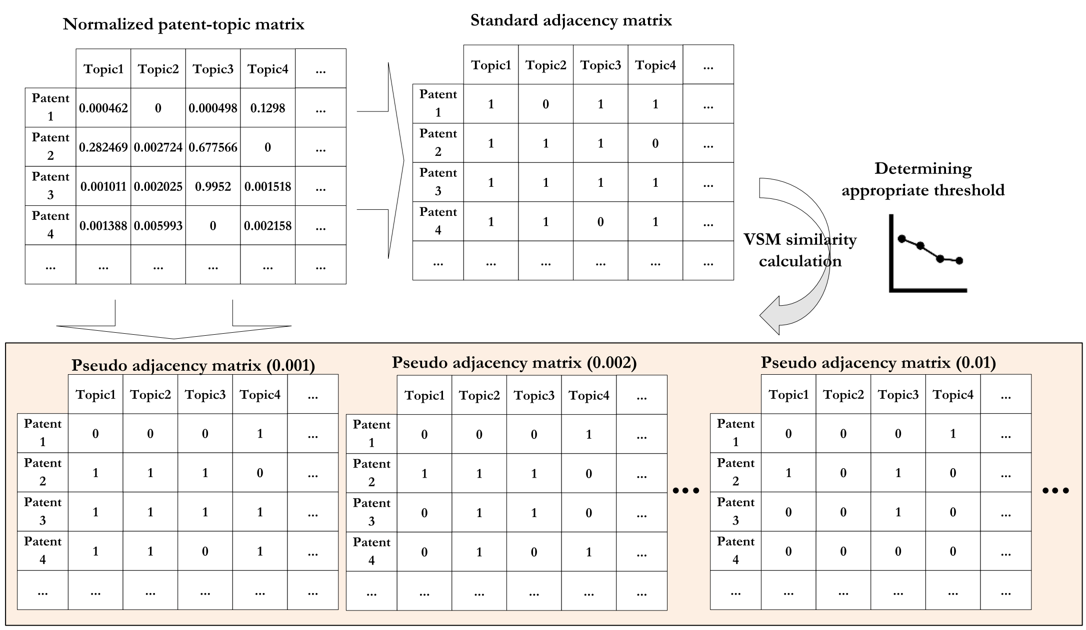



**FIGURE 3.** Process of determining patents' belongingness to innovation topics.

By selecting an appropriate threshold value $\theta$ for each patent and representing its belongingness to innovation topics as a binary value of 0 or 1, a patent-topic adjacency matrix is finally output, which describes the innovation topics to which each patent belongs. This patent-topic adjacency matrix is used as the input for constructing an innovation topic network in the next step.

### *D. CONSTRUCTING INNOVATION TOPIC NETWORKS*

In this step, given the information obtained regarding patents' belongingness to innovation topics, a knowledge network is constructed using the co-occurrences of innovation topics in the same patents. A co-occurrence matrix among various innovation topics in the same patents can then be obtained (Figure 4). The co-occurrence matrix

**Patent-Topic occurrence matrix**

| | Topic1 | Topic2 | Topic3 | Topic4 | Topic5 |
|---|---|---|---|---|---|
| Patent 1 | 0 | 1 | 1 | 0 | 1 |
| Patent 2 | 0 | 0 | 0 | 1 | 1 |
| Patent 3 | 1 | 0 | 1 | 0 | 0 |
| Patent 4 | 0 | 1 | 1 | 0 | 1 |
| Patent 5 | 1 | 1 | 0 | 1 | 0 |

**(Patent-Topic occurrence matrix)$^T$**

| | Patent1 | Patent2 | Patent3 | Patent4 | Patent5 |
|---|---|---|---|---|---|
| Topic 1 | 0 | 0 | 1 | 0 | 1 |
| Topic 2 | 1 | 0 | 0 | 1 | 1 |
| Topic 3 | 1 | 0 | 1 | 1 | 0 |
| Topic 4 | 0 | 1 | 0 | 0 | 1 |
| Topic 5 | 1 | 1 | 0 | 1 | 0 |

**(Patent-Topic) X (Patent-Topic)$^T$**

**Co-occurrence matrix**

| | Topic1 | Topic2 | Topic3 | Topic4 | Topic5 |
|---|---|---|---|---|---|
| Topic 1 | 0 | 1 | 1 | 3 | 1 |
| Topic 2 | 1 | 0 | 0 | 1 | 1 |
| Topic 3 | 1 | 0 | 0 | 1 | 1 |
| Topic 4 | 3 | 1 | 1 | 0 | 1 |
| Topic 5 | 1 | 1 | 1 | 1 | 0 |

Visualization & Computing centrality

**Topic network**

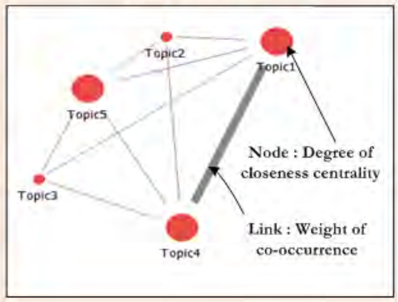



**FIGURE 4.** Process of calculating the co-occurrence of topics and visualizing their relations.

itself represents a network; therefore, an innovation topic network is generated using the matrix. The process of visualizing the network and computing its analysis indices can be supported by various SNA tools, including UCINET (http://www.analytictech.com), Pajek (http://vlado.fmf.uni-lj.si/pub/networks/pajek), and NetMiner (http://www.netminer.com). Previous studies have shown that a well-displayed network visual provides analysts with an overall, intuitive understanding of the entire technology network.

In a friendship network, the most popular actor would be located nearest the center. Similarly, this step adopts closeness centrality to measure an innovation topic's technological importance within the network to which it belongs. A node's closeness centrality in a network is calculated as the sum of the length of the shortest paths between the node and all the others in the network; the more central a node is, the closer it is to all other nodes. The closeness centrality, as the inverse of farness, is defined as follows:

$$C_c(v) = {}^{(N-1)}\Big/_{\sum_{v \neq t \epsilon V} d(v,t)}, \tag{3}$$

where $d(v, t)$ is the distance, which is the shortest path, between nodes $v$ and $t$, and $N$ is the number of nodes in the network for obtaining normalized closeness centrality values.

### *E. GENERATING AN OPPORTUNITY-FOCUSED INNOVATION TOPIC (OFIT) MAP*

The final step involves positioning innovation topics in a two-dimensional space, in order to identify their potential opportunity based on the concept of opportunity algorithm [43]. Some prior studies successfully used the concept of the opportunity algorithm based on importance and satisfaction metrics to create new patents [44] and identify new product opportunities [45]. In this paper, we propose an OFIT map that uses the importance and satisfaction of innovation topics as positioning parameters.

An innovation topic's closeness centrality value is considered as its importance level. From a technological perspective, innovation topics with a high closeness centrality may have strong impact on the technology system; therefore, they can be considered as core or dominant topics in that system. An innovation topic's satisfaction level refers to how much it has accomplished; in this study, this level is estimated based on patent citation. In the literature, it is assumed that the number of patents that cite a former patent can be a proxy for its subsequent development [46]–[49]. In this regard, we consider that a patent contributes a great deal to advancing its relevant technology when it has a high number of forward citations.

Since this approach structures a patent as a vector of invention properties, we can assume that a patent's forward citations are distributed over innovation topics of which it consists. Therefore, each innovation topic's satisfaction level can be evaluated based on its total citation stock over all patents, as follows:

$$CS_i = \sum_{j \in S} CP_{ij} \times FC_{j}, \tag{4}$$

where $CP_{ij}$ is the contribution probability of innovation topic $i$ in patent $j$, $FC_j$ is the number of forward citations of $j$, and $CS_i$ therefore becomes the citation stock of $i$.

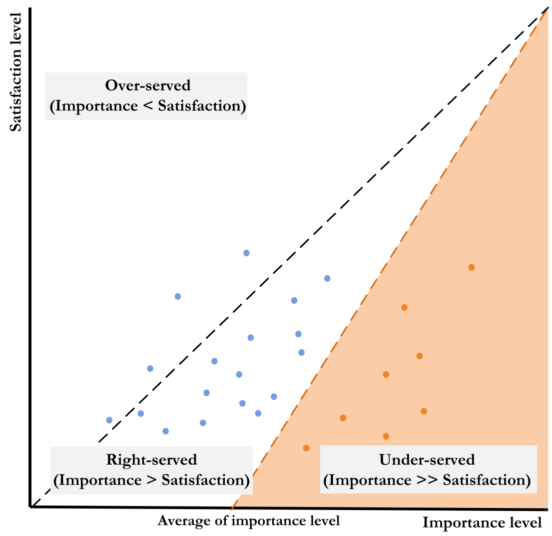



**FIGURE 5.** Schematic of an OFIT map.

According to the opportunity algorithm based on importance and satisfaction [43], each innovation topic can be positioned on our proposed OFIT map, which is divided into three areas: served-right, over-served and under-served (Figure 5). In the opportunity algorithm, the over-served area has relatively high satisfaction compared to its importance, while the under-served area has relatively low satisfaction compared to its importance. Since the over-served innovation topics in the OFIT map are sufficiently developed, they are likely to be original or basic technology. The under-served innovation topics in the OFIT are not technically advanced in comparison to their high importance in the technology network. As a result, these innovation topics can serve as technical opportunities. Furthermore, this OFIT map illustrates a technology system's overall trend, as well as enables experts to identify topics with potential opportunity for directing further R&D.

## IV. CASE STUDY: AUGMENTED REALITY TECHNOLOGY

We use patents relating to augmented reality to illustrate our method. Augmented reality is a live direct or indirect view of a real-world environment integrated with its virtual elements, which transforms the user's environment into digital information in real time by combining virtual objects with the real world. Augmented reality has been increasingly utilized in various fields, such as military, medical, entertainment, and other commercial industries, aided by advanced computing devices such as smartphones and tablet computers. In fact, use of augmented reality-interactive technology provided shoppers with an improved simulated shopping experience and environment [50], [51] and was known to improve consumer decision making [52], [53]. To build an augmented reality system, various technology components, including display, tracker, graphic computer, and software, are required, and need to be combined. That is, augmented reality is not a single technology, but rather a combination of various technical factors. Although augmented reality technology has achieved a rapid growth by its applications, this emerging technology still requires further technological development to be adopted in various industries. Therefore, identifying innovation topics within augmented reality technology and analyzing their trends and potential opportunities would be important for technology planning.

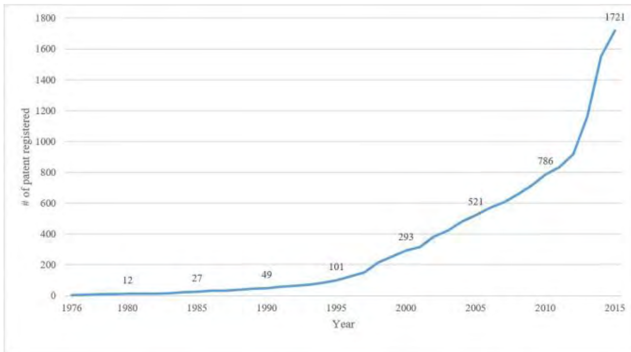



**FIGURE 6.** Yearly patent registration trend.

A total of 1721 patents registered in the Unites States Patent and Trademark Office (USPTO) patent database between 1976 and 2015 were collected using patent search query (Appendix II) as a data source (Figure 6). Overall, prior to the 2000s, patent registrations were fairly constant, with an average annual increase rate of less than 1%, and the cumulative total patent registrations was only 17%. However, since 2000, patent registrations have begun to increase rapidly, with an average increase rate of 5.53%. This increasing tendency has been particularly sharp since 2013, with an average annual increase rate of 15%.

### A. IDENTIFYING AUGMENTED REALITY TECHNICAL INNOVATION TOPICS

The identification of innovation topics is preceded by the extraction of augmented reality keywords. To this end, we used Alchemy API, an open NLP tool, to extract keywords from patent texts. After eliminating irrelevant or too-general words, the valid set consisted of 10,056 keywords, such as "eyepiece," "microcomputer," "virtual camera," "situational location," "radio frequency," and so on. Using this keyword set, an initial patent-keywords matrix was obtained, with the occurrence frequencies of the corresponding keyword in each patent. Then, we generated a weighted patent-keyword matrix by calculating TF-IDF (Table 1).

Next, we applied LDA-based topic modeling to this matrix in order to identify innovation topics within augmented reality technology. To determine the appropriate number of

**TABLE 1.** Portion of patent-keyword matrix with TF-IDF.

| Registration Number | eyepiece | micro computer | virtual camera | Stereoscopic image | game apparatus | optical detector |
|---|---|---|---|---|---|---|
| 8482859 | 1406.49 | 12.63 | 0 | 0 | 0 | 0 |
| 8488246 | 1405.36 | 12.63 | 0 | 0 | 0 | 0 |
| 8869228 | 0 | 900.40 | 0 | 0 | 0 | 0 |
| 8869229 | 0 | 900.40 | 0 | 0 | 0 | 0 |
| 8893177 | 0 | 900.40 | 0 | 0 | 0 | 0 |
| 7870581 | 0 | 893.19 | 0 | 0 | 0 | 0 |
| 8384770 | 0 | 0 | 865.75 | 29.74 | 744.07 | 0 |
| 8060389 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7882222 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8527625 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8814691 | 739.25 | 0 | 0 | 0 | 0 | 0 |
| 8885177 | 0 | 0 | 0 | 0 | 0 | 738.21 |
| 8345952 | 0 | 0 | 0 | 2.83 | 0 | 0 |
| 8749396 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8698772 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8502659 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8427396 | 1.13 | 0 | 0 | 675.55 | 0 | 0 |
| 8894486 | 0 | 0 | 134.67 | 0 | 32.20 | 0 |

topics, we calculated the average cosine similarity between all pairs of topic-keyword distribution vectors (Figure 7). as the number of topics increases, the similarity becomes lower. We found that the average topic similarity degree enters a stabilization phase when the number of topics n is larger than 40. Finally, 57 was selected as the optimal number of topics, because its average topic similarity was the lowest within the range of 40 to 60, where the topic similarities begin to stabilize.

Next, we next obtained two outputs by means of LDA with the 57 topics: a topic-keyword distribution matrix (Table 2) and a patent-topic distribution matrix (Table 3). Finally, 57 innovation topics could be labeled based on the topic-keyword distribution matrix, which illustrates how a topic is composed of its main contributing keywords (Table 4 and Appendix I). For example, topic 43 had keywords such as "gesture," "gesture recognition," "gesture data," and "gesture recognition interface," so this topic was labeled as the "gesture recognition" innovation topic, while topic 14, with keywords "eye," "HUD," "pupil," and "pupil location," was labeled as "Iris scanning." We found that the labeled innovation topics encompassed the majority of augmented reality fields, such as display, tracking, interface, and their applications [54].

The patent-topic distribution matrix illustrates the portion of all patent topics with probability. To determine to

**TABLE 2.** Portion of topic-keyword distribution matrix.

| Topic | eyepiece | micro computer | virtual camera | stereoscopic image | game apparatus | optical detector |
|---|---|---|---|---|---|---|
| Topic-1 | 0.000005 | 0.000005 | 0.000007 | 0.000031 | 0.000005 | 0.000005 |
| Topic-2 | 0.000006 | 0.000006 | 0.000011 | 0.000292 | 0.005872 | 0.000006 |
| Topic-3 | 0.000005 | 0.000007 | 0.000004 | 0.000598 | 0.000004 | 0.000004 |
| Topic-4 | 0.000011 | 0.000004 | 0.000004 | 0.000007 | 0.000004 | 0.026801 |
| Topic-5 | 0.000005 | 0.000005 | 0.000005 | 0.000005 | 0.000005 | 0.000005 |
| Topic-6 | 0.000004 | 0.000004 | 0.000005 | 0.000004 | 0.000005 | 0.000004 |
| Topic-7 | 0.000005 | 0.000005 | 0.000005 | 0.000006 | 0.000005 | 0.000005 |
| Topic-8 | 0.000005 | 0.000005 | 0.000009 | 0.000005 | 0.000005 | 0.000005 |
| Topic-9 | 0.000014 | 0.000006 | 0.000005 | 0.000018 | 0.000028 | 0.000005 |
| Topic-10 | 0.000003 | 0.000003 | 0.000003 | 0.000003 | 0.000003 | 0.000006 |

**TABLE 3.** Portion of patent-topic distribution matrix.

| Registration no. | Topic-1 | Topic-2 | Topic-3 | Topic-4 | Topic-5 | Topic-6 |
|---|---|---|---|---|---|---|
| 3940204 | 0.002098 | 0.001822 | 0.002121 | 0.085846 | 0.001891 | 0.002947 |
| 3945716 | 0.034489 | 0.007274 | 0.072926 | 0.007804 | 0.007185 | 0.035196 |
| 4001499 | 0.005503 | 0.005895 | 0.391804 | 0.005699 | 0.007526 | 0.006743 |
| 4026641 | 0.008331 | 0.009243 | 0.007875 | 0.03716 | 0.007328 | 0.007601 |
| 4028725 | 0.066186 | 0.003958 | 0.011743 | 0.004058 | 0.004008 | 0.004707 |
| 4152846 | 0.003214 | 0.002008 | 0.002419 | 0.002521 | 0.002008 | 0.002162 |
| 4153913 | 0.003202 | 0.002222 | 0.005247 | 0.002866 | 0.002222 | 0.003734 |
| 4303394 | 0.004613 | 0.001829 | 0.242306 | 0.001575 | 0.001653 | 0.001947 |
| 4303868 | 0.004905 | 0.004556 | 0.018749 | 0.004788 | 0.004614 | 0.004614 |

which innovation topics a patent belongs, the optimized cut-off value was determined based on the similarity calculation between a standard adjacency VSM and its pseudo adjacency VSMs, using Eq. 2. As a result, 95% of the maximum match was found to have 0.038 (Figure 8). Using this optimal threshold value of 0.038, a patent-topic

**TABLE 4.** Portion of innovation topic labeling results.

| Topic # | Innovation topic | Highly contributing keywords and topic description |
|---|---|---|
| Topic-2 | Panoramic image | Panoramic image, underground facility, aerial image, display controller, street view, image data, foreground image |
| | | Innovation topic related to image for the wide-angle view or whole aerial image. |
| Topic-4 | Light modulator | Light modulator, spatial light, prism, optics, optical detector, lens, reflective surfaces |
| | | Innovation topic related to optical setup for adjusting the light to make an optical image |
| Topic-5 | Feature extraction | Capture device, image patch, feature extraction, matching module, media server, |
| | | Innovation topic related to image analysis and processing for setting certain features and extracting information from image and e |
| Topic-14 | Iris scanning | Eye, HUD, pupil, pupil location, normalized pupil, gaze axis, light sources, motion capture |
| | | Innovation topic related to tracking and recognition with scanning the user's eyes |
| Topic-17 | Photographing apparatus | Imaging apparatus, photographed image, photographing device, image processing, stereo image, photographing position |
| | | Innovation topic related to image apparatus and processing with the photographing device and its image |
| Topic-18 | Image capture based on marker | Maker, image capturing, image coordinate, capture device, orientation sensor, marker detecting, orientation measurement |
| | | Innovation topic related to tracking the movement and position of objects using marker |
| Topic-39 | Display panel | Liquid crystal panel, insulating film, projection surface, interlayer insulating film, light pattern, |
| | | Innovation topic for display screen panel |
| Topic-43 | Gesture recognition | Gesture, gesture recognition, mobile access, eyesight, gesture data, input gesture, gesture set |
| | | Innovation topic for recognizing user's gesture and processing it |
| Topic-48 | Mobile location sensing | Object space, portable computing, portable computing device, magnetic field detector, geographic location, map information, mobile phone, |
| | | Innovation topic for location sense used with portable computing device |
| Topic-51 | Space image sensing | Space image, image sensing, virtual space, image processing apparatus, real space, virtual spaced image, |
| | | Innovation topic for image processing for space image |

adjacency matrix was generated (Table 5); this matrix represents the innovation topics to which each patent belongs. For example, patent no. 3940204, "optical display systems utilizing holographic lenses" contains four topics: topic 4 (Light modulator), topic 30 (Video display screen for HMD), topic 44 (Optical image processing), and topic 56 (Portable device for augmented reality image (HW)). Patent no. 3940204, "speed sensor and head-mounted data display for sportsman or skier" contains two topics: topic 2 (Panoramic image) and topic 25 (Context awareness). The average number of topics per patent was found to be 3.826.

**TABLE 5.** Portion of patent-topic adjacency matrix.

| Patent no | Topic-1 | Topic-2 | Topic-3 | Topic-4 | Topic-5 | Topic-6 | Topic-7 |
|---|---|---|---|---|---|---|---|
| 3940204 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 3945716 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| 4001499 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 4026641 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 4028725 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 4152846 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4153913 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4303394 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 4303868 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

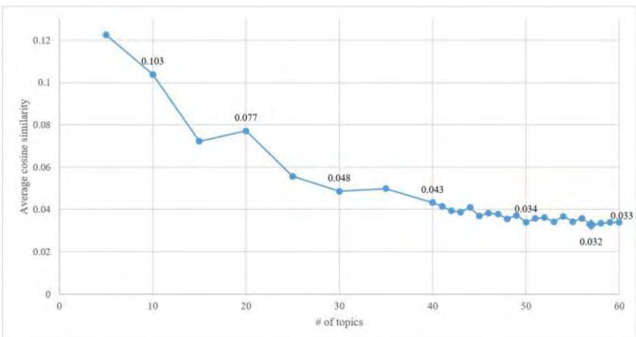



**FIGURE 7.** Average cosine similarity according to number of topics.

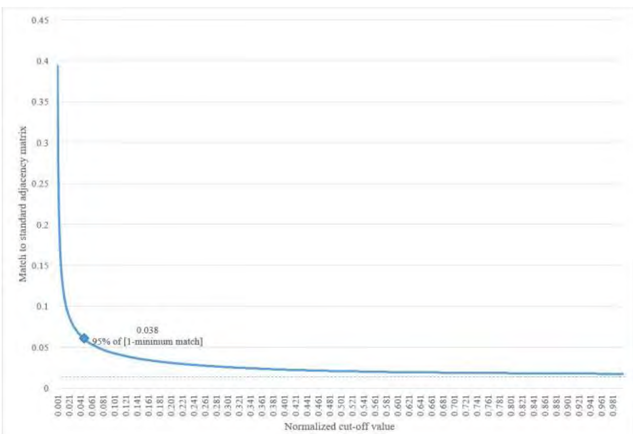



**FIGURE 8.** Similarities between standard and pseudo adjacency matrices by cut-off value.

## *B. GENERATING INNOVATION TOPIC NETWORKS AND OFIT MAPS*

Based on the patent-topic adjacency matrix, a co-occurrence matrix of topics within the same patents was obtained using Eq. 2 (Table 6). The co-occurrence matrix itself is a network, and illustrates the relationships between innovation topics.

In this study, we employed a commercial SNA tool, NetMiner, to visualize the innovation topic network and compute the topics' centrality values (Figure 9). The network consists of 57 innovation topics and their relations. The node size is the closeness centrality of each innovation topic, while the link width is the relationship strength between the two concepts. To enable intuitive understanding of our network, innovation topics were grouped together into five major fields: hardware, display, mobile device, tracking, and application. For example, the hardware group, which represents the mechanical body for realizing augmented reality, includes innovation topics such as wearable computing device, display unit (HW), and display panel. The application group includes cases applied to real life, such as applications for game apparatuses, vehicles, and commerce. Therefore, a visualized network provides an effective representation and supports an overall, intuitive understating of innovation topics and their relationships. However, for in-depth network analysis and quantification, the closeness centrality needs to be calculated for deriving innovation topics that play a major role.

From the network, the closeness centrality of each innovation topic was calculated, using Eq. 3 as the importance level (Table 7). Innovation topics that are highly ranked in terms of centrality have a strong relationship with other topics, which means they are dominant or core concepts in augmented reality. Therefore, closeness centrality values are used as a proxy for representing innovation topics' level of importance.

The average closeness centrality value for all innovation topics was 0.5012. According to the closeness centrality analysis, the innovation topic with the highest importance value was topic 3, ''Processing video signal'' (0.79438), followed by topic 33, ''Image generator for HMD'' (0.666887) and topic 44, ''Optical image processing'' (0.650818). These top three innovation topics reflect that technology related to image processing for display is the core and dominant innovation topic in augmented reality systems. In addition, topic 23, ''Mobile location information'' and topic 45,

**TABLE 6.** Portion of co-occurrences among innovation topics.

| Topic | Topic-1 | Topic-2 | Topic-3 | Topic-4 | Topic-5 | Topic-6 | Topic-7 |
|---|---|---|---|---|---|---|---|
| Topic-1 | 0 | 6 | 34 | 9 | 9 | 33 | 4 |
| Topic-2 | 6 | 0 | 19 | 4 | 4 | 11 | 5 |
| Topic-3 | 34 | 19 | 0 | 25 | 12 | 29 | 8 |
| Topic-4 | 9 | 4 | 25 | 0 | 2 | 8 | 7 |
| Topic-5 | 9 | 4 | 12 | 2 | 0 | 10 | 3 |
| Topic-6 | 33 | 11 | 29 | 8 | 10 | 0 | 2 |
| Topic-7 | 4 | 5 | 8 | 7 | 3 | 2 | 0 |
| Topic-8 | 15 | 5 | 21 | 8 | 6 | 19 | 3 |
| Topic-9 | 20 | 10 | 50 | 26 | 4 | 18 | 9 |

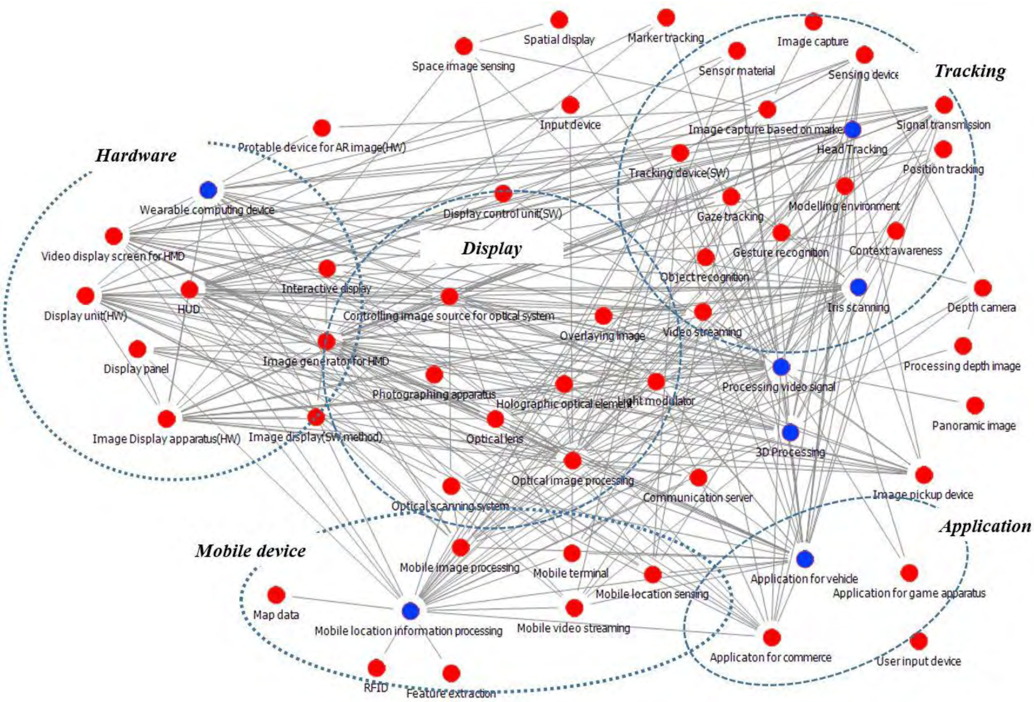



**FIGURE 9.** Network of innovation topics.

''Mobile image processing'' were ranked within the top level, which means that compact and portable computing processing and devices such as smart phones are also important concepts in augmented reality. The seventh-ranked topic 14, ''Iris scanning'' (0.628115) and eighth-ranked topic 38, ''Head tracking'' (0.606942) relate to recognizing human body part movements, which means that detecting human body parts, especially the eyes and head, as well as their interaction and movement in relation to augmented reality devices, is important.

The satisfaction levels of innovation topics were computed using Eq. 4 (Table 8). Innovation topics with a high citation stock can be considered as contributing to the advancement of the relevant technology. The average satisfaction value was 951.3158. Topic 44, ''Optical image processing'' had the highest value with 3215.717, followed by topic 33, ''Image generator for HMD,'' with 2524.688 and topic 3, ''Processing video signal,'' with 2355.084. These top three innovation topics relate to creating and processing image sources for visual displays. In fact, among the various sensory displays,

**TABLE 7.** Portion of closeness centrality values of innovation topics.

| Rank | Topic | Innovation topic | closeness |
|---|---|---|---|
| 1 | Topic-3 | Processing video signal | 0.79438 |
| 2 | Topic-33 | Image generator for HMD | 0.666887 |
| 3 | Topic-44 | Optical image processing | 0.650818 |
| 4 | Topic-23 | Mobile location information processing | 0.64307 |
| 4 | Topic-28 | Application for vehicle | 0.64307 |
| 6 | Topic-54 | HUD | 0.635504 |
| 7 | Topic-14 | Iris scanning | 0.628115 |
| 8 | Topic-38 | Head Tracking | 0.606942 |
| 9 | Topic-6 | 3D Processing | 0.593603 |
| 10 | Topic-10 | Wearable computing device | 0.587151 |
| 10 | Topic-30 | Video display screen for HMD | 0.587151 |
| 10 | Topic-9 | Video streaming | 0.587151 |
| 13 | Topic-1 | Tracking device(SW) | 0.580837 |
| 14 | Topic-26 | Controlling image source for optical system | 0.551203 |
| 14 | Topic-34 | Image Display apparatus(HW) | 0.551203 |
| 14 | Topic-45 | Mobile image processing | 0.551203 |
| 14 | Topic-55 | Signal transmission | 0.551203 |
| 18 | Topic-15 | Sensing device | 0.540179 |
| 18 | Topic-19 | Display unit(HW) | 0.540179 |
| 18 | Topic-53 | Optical lens | 0.540179 |

such as visual, aural, and olfactory displays, that of visual is much further developed and currently practical. However, other displays are less developed or do not yet exist as augmented reality displays. Topic 45, "Mobile image processing," which ranked fifth, and topic 23, "Mobile location information processing," which ranked ninth, relate to the mobile unit for displaying images or computing information. In addition, topic 1, "Tracking device (SW)" was ranked sixth; this innovation topic is strongly related to processing data from tracking or recognition.

Next, we generated an augmented reality OFIT map based on two parameters: importance and satisfaction levels (Figure 10). To this end, we first normalized both values using the maximum and minimum of each to adjust the range of each axis. By building the normalized average of the importance, 0.406117, the origin (0, 0) and the maximum value (1, 1), the map was divided into three sections: the over-served, served-well and under-served areas.

Overall, the majority of innovation topics are located in the right-served area, which means that these concepts are developed well, with balanced levels of importance and satisfaction. For example, topic 43, "Gesture recognition," was located in the over-served area on the OFIT map. This is the innovation topic relating to interpreting a user's actions, such as hand gestures or states, by means of mathematical algorithms, and could constitute input data and one of the techniques for user interaction. US patent 7701439 is a gesture recognition simulation system method; this topic has recently been relatively well developed, and we found that input gestures and gesture data for mobile access have already been commercialized in the mobile phone and tablet computer market.

However, innovation topics 3, "Processing video signal," 28, "Application for Vehicle," 23, "Mobile location information processing," 14, "Iris scanning," 38, "Head Tracking," and 10, "Wearable computing device," are in the under-served area, where innovation topics have a low level of satisfaction compared to their importance.

Although the topic 3, "Processing video signal," is located in the under-served area, its importance and satisfaction levels were both high. In fact, this innovation topic formed links with 43 others in the network, and ranked the highest in terms of closeness centrality, as well as having the third highest citation stock among the 57 innovation topics. In fact, the use

**TABLE 8.** Portion of innovation topic citation stocks

| Rank | Topic | Innovation topic | Stock of citation |
|---|---|---|---|
| 1 | Topic-44 | Optical image processing | 3215.717 |
| 2 | Topic-33 | Image generator for HMD | 2524.688 |
| 3 | Topic-3 | Processing video signal | 2355.084 |
| 4 | Topic-34 | Image Display apparatus(HW) | 2326.809 |
| 5 | Topic-45 | Mobile image processing | 1771.705 |
| 6 | Topic-1 | Tracking device(SW) | 1676.051 |
| 7 | Topic-54 | HUD | 1582.406 |
| 8 | Topic-28 | Application for vehicle | 1315.767 |
| 9 | Topic-23 | Mobile location information processing | 1228.393 |
| 10 | Topic-20 | Image pickup device | 1212.816 |
| 11 | Topic-43 | Gesture recognition | 1187.418 |
| 12 | Topic-30 | Video display screen for HMD | 1175.702 |
| 13 | Topic-26 | Controlling image source for optical system | 1130.672 |
| 14 | Topic-50 | Depth camera | 1086.872 |
| 15 | Topic-39 | Display panel | 1075.34 |
| 16 | Topic-9 | Video streaming | 1073.976 |
| 17 | Topic-25 | Context awareness | 1070.483 |
| 18 | Topic-18 | Image capture based on marker | 1064.367 |
| 19 | Topic-48 | Mobile location sensing | 1036.506 |
| 20 | Topic-4 | Light modulator | 1035.471 |

of video images in display is a traditional approach; therefore, video signal processing has been important and steadily well developed. For example, US patent 4786966 relates to a video system for a head-mounted display, and a method for transmitting video signals to a remotely located video display. It was registered in 1988 and intended for military and weapons applications. US patent 6578203, registered in 2003, relates to a video signal distribution system for a head-mounted display, and includes an interface device for receiving a plurality of video signals from various sources, combing these signals into various forms, and transmitting the combinations to a receiver. Furthermore, this method uses a portable device, including a display and receiver, and transmits signals via wireless RF. Therefore, topic 3, ''Processing video signal'' has been used in a wide range applications, including military, space, and recently, entertainment display devices, with portable, miniaturization, and wireless features. As a result, it is a major, original innovation topic and can serve as the basis for facilitating other adjacent innovation topics, such as topics 13, ''Application for game apparatus,'' and 45, ''Mobile image processing.''

However, all of the innovation topics, except for topic 3, in the under-served area do not have a high level of satisfaction with respect to their importance; that is, these topics need to be furthered. Topic 38, ''Head tracking,'' is an important concept for tracking head movements. Most augmented reality display technologies are based on head-mounted displays. Previously, there was a function for simply displaying images through HMD; however, as is the case in many fields, user interaction has become increasingly important, and the user's head movements need to be tracked and reflected. US patent 5742263, registered in 1998, relates to an HMD system that has a processor to control the video information depicted for head tracking purposes. Furthermore, US patent 8754931, registered in 2014, relates to video eyewear for smartphone games, and provides a method for improving user experience with connections, software programming, and interaction between a smartphone and HMD or other video eyewear. To enable interaction, this method makes use of head tracking as an input parameter for display, to change the user's viewpoint in a game apparatus or ''smart'' HMD. This topic was linked to a total of 22 others in the network, including topic 1,

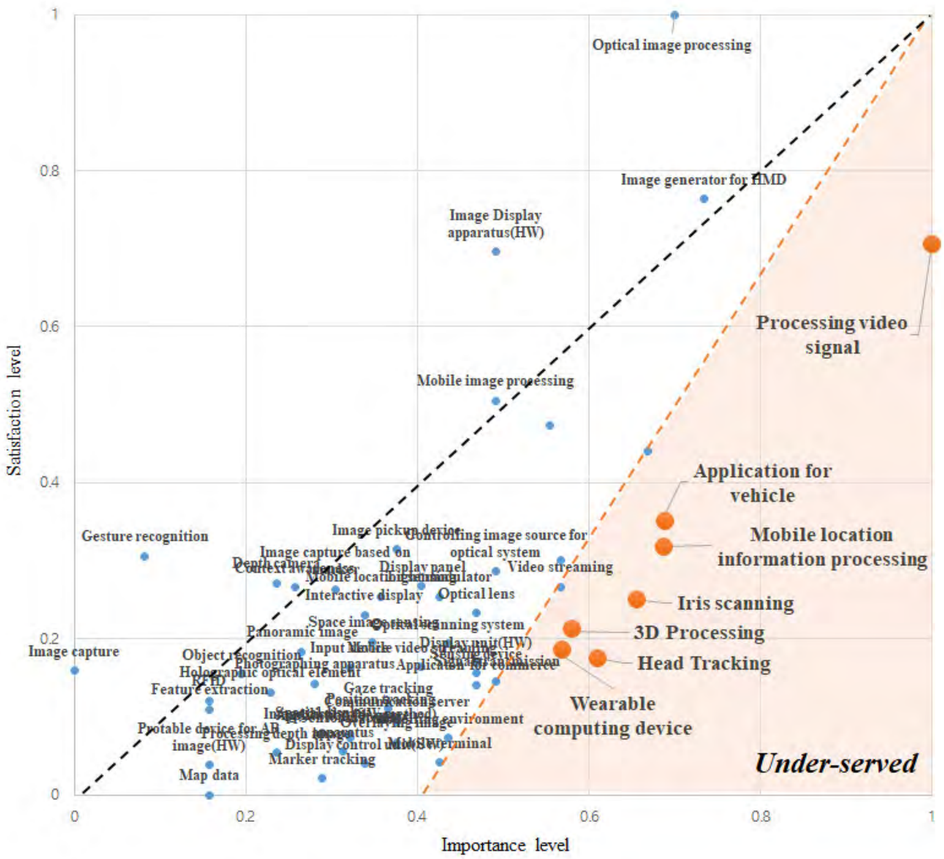



**FIGURE 10.** **Augmented reality OFIT map.**

"Tracking device (SW)," topic 10, "Wearable computing device," and topic 14, "Sensing device." Although it is a widely used factor, it is still underdeveloped; thus, we consider that it has potential opportunity for improving tracking and display capabilities.

Topic 28, "Application for vehicle (tracking)" is an innovation topic relating to display in the automotive field. It is used in the automobile industry to display information to drivers, such as the vehicle status and a navigation map; therefore, sensor, image display, and location information technologies were applied. US patent 4504910 deals with current position display, adapted for mounting in a vehicle. US patent 5142274 relates to a method for a silhouette-illuminated vehicle head-up display apparatus, while US patent 9008369 relates to a vehicle vision system, including an operable camera to capture image data for various driver assistance systems. US patent 8761962 deals with a system for controlling an in-vehicle device using augmented reality, and includes a mobile device to identify a vehicle object as a unit and receive control commands. US patent 5784036 relates to a head-up display with a selective function for enhanced driver recognition. While many technical elements have been applied and utilized, development of the related technologies has been insufficient. At present, with smart car technology such as self-driving cars, augmented reality technology can provide the opportunity to enhance the driver's safety and convenience.

Topic 23, "Mobile location information processing," was also found to be under-served, and its mobile device with the mobility feature can further realize location-based services. To achieve this, the mobile device communicates with the server to handle geographic location information. This topic was linked to a total of 27 others in the network, including topic 10, "Wearable computing device," topic 46, "RFID," and topic 7, "Map data." For example, US patent 8239132, registered in 2012, deals with a method for delivering location-oriented information, including a location server and mobile location device, which contains a geographical positioning module to determine location information. US patent 8442502, registered in 2013, relates to tracking technologies in terms of an object's location in augmented reality, using a mobile phone. At present, mobile devices such as smartphones and tablet PCs are being utilized as augmented reality hardware. Therefore, it is important for mobile devices to process location information according to the user's movement, and provide virtual information. Despite these factors, this innovation topic is not yet sufficiently developed, which means that an opportunity exists.

Moreover, opportunity is revealed in terms of hardware that can realize augmented reality technology, in topic 10, "Wearable computing device." Although hardware apparatuses have been considered and associated with other innovation topics, its degree of development is lower than its importance. Thus far, the representative apparatus for augmented reality display has been HMD. US patent 5299063, registered in 1994, relates to a cross projection visor helmet mounted display, while US patent 5138555, registered in 1992, relates to helmet mounted display adaptive predictive tracking. However, in recent years, various patents regarding wearable devices have been registered: US patent 7249846, registered in 2007, relates to eyewear with an image, while US patent 9024843, registered in 2015, outlines a curved display for a wearable computer. Furthermore, in our innovation topic network, this topic has relations to topic 12, "Interactive display," and topic 38, "Head tracking." Since this concept is an opportunity, R&D needs to be applied to various wearable devices, such as smart watches and Google Glass, in order to provide users with the value of light and easy apparatus.

Topic 14, "Iris Scanning," is an innovation topic for recognizing human eye movements. In augmented reality, analyzing the user's eye movements provides useful information for display and interaction; thus, this topic can relate to display apparatus or signal processing. US patent 5583795 deals with an apparatus for measuring eye gazes, while US patent 6396461 relates to personal display with vision tracking, including an eye position detector to monitor the light reflected from the user's eye to identify the pupil position. Furthermore, US patent 6496461 relates to a head-mounted display with an eye-tracking capability, based on the reflection of LEDs at the cornea of a user's eye. However, eye movement is a microscopic change, unlike that of the head. Therefore, nanoscale changes need to be detected using biosensors. For this reason, the technology is still not sufficiently developed. If a technology could relate to this topic, it would provide an opportunity for offering a hands-free operation system.

## V. DISCUSSION AND CONCLUDING REMARKS

An innovation topic is a generalized technical subject that there is a desire to develop or improve within a technology system. Since patents are technical documents regarding inventions, they are effective materials for identifying innovation topics. However, prior patent-based studies have analyzed technology trends in terms of only patents or keywords, rather than using latent innovation topics. As an alternative, this study proposes a method for identifying innovation topics from patents and thereby generating an OFIT map. The proposed method is based on LDA-based topic modeling, SNA, and an opportunity algorithm. In particular, it enables experts to identify useful directions for R&D strategies and business planning.

**TABLE 9. List of 57 labeled topics**

| # of Topics | Topic Label |
|---|---|
| Topic-1 | Tracking device(SW) |
| Topic-2 | Panoramic image |
| Topic-3 | Processing video signal |
| Topic-4 | Light modulator |
| Topic-5 | Feature extraction |
| Topic-6 | Projective display |
| Topic-7 | Map data |
| Topic-8 | Modelling environment |
| Topic-9 | Video streaming |
| Topic-10 | HMD(HW) |
| Topic-11 | Sensor material |
| Topic-12 | Interactive display |
| Topic-13 | Application for game apparatus |
| Topic-14 | HUD(head up display) |
| Topic-15 | Sensing device |
| Topic-16 | Input device |
| Topic-17 | Photographing apparatus |
| Topic-18 | Marker |
| Topic-19 | Display unit(HW) |
| Topic-20 | Image pickup device |
| Topic-21 | Mobile video streaming device |
| Topic-22 | Position tracking |
| Topic-23 | Mobile location information processing |
| Topic-24 | Image display(SW method) |
| Topic-25 | Context awareness |
| Topic-26 | Controlling image source for optical system |
| Topic-27 | Communication server |
| Topic-28 | Application for vehicle(image processor) |
| Topic-29 | Display control unit(SW) |
| Topic-30 | HMD(video display) |
| Topic-31 | Optical scanning system |
| Topic-32 | Object recognition |
| Topic-33 | HMD(optical display) |
| Topic-34 | Optical image display apparatus(HW) |
| Topic-35 | Identification of marker object |
| Topic-36 | Processing depth image |
| Topic-37 | Holographic optical element |
| Topic-38 | Computer generated image |
| Topic-39 | Display panel |
| Topic-40 | Remote Controller |
| Topic-41 | Capture device |
| Topic-42 | Overlaying image |
| Topic-43 | Gesture recognition |
| Topic-44 | Optical image processing |
| Topic-45 | Mobile image processing |
| Topic-46 | RFID |
| Topic-47 | Spatial display |

**TABLE 9.** ***(Continued.)*** **List of 57 labeled topics**

| | |
|---|---|
| Topic-48 | Portable computing device |
| Topic-49 | User input device |
| Topic-50 | Depth camera |
| Topic-51 | Space image sensing |
| Topic-52 | Gaze tracking |
| Topic-53 | Optical lens |
| Topic-54 | Application for vehicle(tracking) |
| Topic-55 | Signal transmission |
| Topic-56 | Portable device for AR image(HW) |
| Topic-57 | Application for commerce |

In this paper, we have presented a procedure for generating an OFIT map. Firstly, keywords are extracted from the collected and preprocessed patent documents for a specific technology, and each patent is structured as a keywords vector using VSM. Innovation topics are identified according to LDA topic modeling, and their relationships are visualized by means of an innovation topic network. Next, each innovation topic is evaluated according to two parameters: its importance and satisfaction levels. The importance level is computed by the closeness centrality centrality in the network, while the satisfaction level is calculated using stock of citations. The OFIT map is generated based on these two parameters, and from the map, innovation topics that have a high technical importance but low satisfaction, and are therefore underserved, can be identified as opportunities for developing the given technical field.

This paper has illustrated the proposed methodology for augmented reality technology, using 1,721 US patents from 1976 to 2015. From the patents, a total of 57 topics were extracted as innovation topics that describe the technical subjects in augmented reality. We evaluated each innovation topic in terms of two factors, namely importance and satisfaction. Finally, the OFIT map was generated, and various underserved innovation topics were identified: ''Head tracking'' for interaction with users, ''Application for vehicle'' for displaying information to drivers, ''Mobile location information processing'' with the mobility feature to realize location-based services, ''Wearable computing device'' for various types of display devices, and ''Iris scanning'' to recognize human eye movements. These innovation topics have opportunity potential because they were not sufficiently developed in relation to their importance. According to our analysis, the innovation topic ''Wearable computing device'' provides an opportunity, as various wearable devices such as the smart watch and Google Glass have recently been developed. Therefore, this technology can be applied to provide users with a light and convenient display device.

It is expected that our study will provide contributions to both academia and industry to relevant fields. First, methodologically, our proposed method provides a different type of view on technology analysis, by focusing on technical topics or subject matters that the technology in question is attempting to develop or improve. While most prior approaches addressed information mapping and its interpretation building on the relatedness of patents or keywords, this study proposed a procedure to analyze the relationship among innovation topics and evaluate them in terms of opportunity potential. Second, our approach has the potential to help generate insights into various technology domains. This is because this approach is neutral with respect to the type of data used, although we utilized patent data relating to augmented reality technology. Based on patent data in any technology domain R&D project managers or researchers can apply our approach to identify the relationship among technical innovation topics and their potential, thereby

**TABLE 10.** **Patent search query**

| Patent Retrieval Query in WIPSON |
|---|
| (((augmented near2 (reality or experience)) or (mixed near2 reality)))OR(((augmented near2 (reality or experience)) or (mixed near2 reality)) and (track* or trace* or chase or pursuit* or pursue) and (Gyro* or INS or (inertia adj2 navi-gation) or sensor sens* or gis or GPS or (global near2 positioning) or LBS or (location adj2 based adj2 service) or (positonal adj3 information) or compass or ((magnetic or electromagnetic or mechanical or optical or inertial or ultrasonic) near2 tracker) or track*))OR(((augmented near2 (reality or experience)) or (mixed near2 reality)) and (marker or ARTag or Artoolkit or colorcode or QR-code or ((identification or specific or AR) near2 code)))OR(((augmented near2 (reality or experience)) or (mixed near2 reality)) and (track* or trace* or chase or pursuit* or pursue) and (markerless or SIFT or ferns or ((edge or Feature) near3 extract*) or slam or (interest near2 point) or POI or camera or ((3d or 3-dimension*) adj2 (model or based)) or (computer near2 vision)))OR(((augmented near2 (reality or experience)) or (mixed near2 reality)) and (track* or trace* or chase or pursuit* or pursue) and (((Gyro* or INS or (inertia adj2 navigation) or gis or GPS or (global near2 positioning) or LBS or (location adj2 based adj2 service) or (positonal adj3 information) or compass or ((magnetic or electromagnetic or mechanical or optical or inertial or ultrasonic) near2 tracker)) and ((marker or ARTag or Artoolkit or colorcode or QR-code or ((identification or specific or AR) near2 code)) or (markerless or SIFT or ferns or ((edge or Feature) near3 extract*) or slam or (interest near2 point) or POI or camera or ((3d or 3-dimension*) adj2 (model or based)) or (computer near2 vision)))) or ((marker or ARTag or Artoolkit or colorcode or QR-code or ((identification or specific or AR) near2 code)) and (markerless or SIFT or ferns or ((edge or Feature) near3 extract*) or slam or (interest near2 point) or POI or camera or ((3d or 3-dimension*) adj2 (model or based)) or (computer near2 vision)))))OR((((augmented near2 (reality or experience)) or (mixed near2 reality)) and (display or express or show or project* or screen) and (HMD OR HUD OR EMD OR parallax OR ((head-mount*)or(head and mount*)) OR head-up OR Eyeglass-Mounted or ((video or optical) near2 see-through))) OR(((head and mount*) or (head-mount*)) and dis-play))OR(((augmented near2 (reality or experience)) or (mixed near2 reality)) and (display or express or show or project* or screen) and (project* OR Spatial OR Diorama OR SAR or non-HMD or spatial-AR or laser or calibration or hologra* or panorama or cineorama))OR(((augmented near2 (reality or experience)) or (mixed near2 reality)) and (display or express or show or project* or screen) and (Smartphone OR UMPC OR tablet OR PDA OR handheld OR mobile OR ((personal or portable) near2 terminal) OR (personal adj2 digital adj2 assistant) or (mobile adj2 communication and device))) |

developing further R&D plans. From an industrial perspective, the results of this study can provide decisive information for experts who are working in the field. For example, the innovation topics and "Application for vehicle" and "Wearable computing device" had much opportunity with respect to industry application. Next, our approach can be performed computationally. Because recent technologies change rapidly and are more dynamic than ever before, it is important for firms to track and deal with rapidly evolving technological trends to retain their technical competitiveness. Therefore, our methodology will be an effective tool for monitoring new or emerging technology, by identifying key innovation topics with the possibility of opportunity. Moreover, as this study proposes quantified analysis, such as determining thresholds, it may reduce the intuition or subjective judgment of a person, compared to existing approaches. Thus, our method may provide useful tools for analyzing trends in the specific technology by implementing an automated system.

However, this research still exhibits certain limitations. In our proposed method, we conducted a static analysis only during one period. In order to provide an understanding of the dynamic trends of the innovation topic over time, further research should separate the period into several intervals, and then identify the panoramic trends of the augmented reality technology. Furthermore, although a patent is an effective source, it does not consider market information; thus, in further research, we can evaluate innovation topics by means of data sources that consider market information. Finally, this research has only attempted to identify innovation topics related to augmented reality. In the future, new and diverse technology will emerge; therefore, if it is applied to other areas, our methodology can provide meaningful insight for technology trend analysis.

## APPENDIX I
See Table 9.

## APPENDIX II
See Table 10.

## REFERENCES

[1] S. Maranville, "Entrepreneurship in the business curriculum," *J. Edu. Bus.*, vol. 68, no. 1, pp. 27–31, 1992.

[2] M. J. C. Martin, *Managing Innovation and Entrepreneurship in Technology-Based Firms*, vol. 20. New York, NY, USA: Wiley, 1994.

[3] J. H. Mikkola, "Portfolio management of R&D projects: Implications for innovation management," *Technovation*, vol. 21, no. 7, pp. 423–435, 2001.

[4] N. O'Regan and A. Ghobadian, "The importance of capabilities for strategic direction and performance," *Manage. Decision*, vol. 42, no. 2, pp. 292–313, 2004.

[5] X. Jin, "Approach to IT innovation concept network based on analysis of discourse," in *Advances in Electric and Electronics*. Berlin, Germany: Springer, 2012, pp. 205–212.

[6] J. Yoon, M. Kim, D. Kim, and J. Kim, "Monitoring the change of technological impacts of technology sectors using patent information," *Ind. Eng. Manage. Syst.*, vol. 14, no. 1, pp. 58–72, 2015.

[7] M. Kim, H. Park, and J. Yoon, "Generating patent development maps for technology monitoring using semantic patent-topic analysis," *Comput. Ind. Eng.*, vol. 98, pp. 289–299, Aug. 2016.

[8] H. Park, J. Yoon, and K. Kim, "Identification and evaluation of corporations for merger and acquisition strategies using patent information and text mining," *Scientometrics*, vol. 97, no. 3, pp. 883–909, 2013.

[9] B. U. Yoon, C. B. Yoon, and Y. T. Park, "On the development and application of a self–organizing feature map–based patent map," *R&D Manage.*, vol. 32, no. 4, pp. 291–300, 2002.

[10] J. Yoon and K. Kim, "Detecting signals of new technological opportunities using semantic patent analysis and outlier detection," *Scientometrics*, vol. 90, no. 2, pp. 445–461, 2011.

[11] B. Yoon and Y. Park, "A text-mining-based patent network: Analytical tool for high-technology trend," *J. High Technol. Manage. Res.*, vol. 15, no. 1, pp. 37–50, 2004.

[12] J. Yoon and K. Kim, "Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks," *Scientometrics*, vol. 88, no. 1, pp. 213–228, 2011.

[13] P.-L. Chang, C.-C. Wu, and H.-J. Leu, "Using patent analyses to monitor the technological trends in an emerging field of technology: A case of carbon nanotube field emission display," *Scientometrics*, vol. 82, no. 1, pp. 5–19, 2010.

[14] C. Lee *et al.*, "Monitoring trends of technological changes based on the dynamic patent lattice: A modified formal concept analysis approach," *Technol. Forecast. Soc. Change*, vol. 78, no. 4, pp. 690–702, 2011.

[15] J. Choi and Y.-S. Hwang, "Patent keyword network analysis for improving technology development efficiency," *Technol. Forecast. Soc. Change*, vol. 83, pp. 170–182, Mar. 2014.

[16] J. Yoon, S. Choi, and K. Kim, "Invention property-function network analysis of patents: A case of silicon-based thin film solar cells," *Scientometrics*, vol. 86, no. 3, pp. 687–703, 2010.

[17] H. Park, K. Kim, S. Choi, and J. Yoon, "A patent intelligence system for strategic technology planning," *Expert Syst. Appl.*, vol. 40, no. 7, pp. 2373–2390, 2013.

[18] M.-Y. Wang, D. S. Chang, and C. H. Kao, "Identifying technology trends for R&D planning using TRIZ and text mining," *R&D Manage.*, vol. 40, no. 5, pp. 491–509, 2010.

[19] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Commun. ACM*, vol. 18, no. 11, pp. 613–620, 1975.

[20] P. D. Turney and P. Pantel, "From frequency to meaning: Vector space models of semantics," *J. Artif. Intell. Res.*, vol. 37, pp. 141–188, Feb. 2010.

[21] L. Jing, M. K. Ng, and J. Z. Huang, "Knowledge-based vector space model for text clustering," *Knowl. Inf. Syst.*, vol. 25, no. 1, pp. 35–55, 2010.

[22] J. Leskovec, *Mining of Massive Datasets*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[23] S. Robertson, "Understanding inverse document frequency: On theoretical arguments for IDF," *J. Documentation*, vol. 60, no. 5, pp. 503–520, 2004.

[24] C. Son, Y. Suh, J. Jeon, and Y. Park, "Development of a GTM-based patent map for identifying patent vacuums," *Expert Syst. Appl.*, vol. 39, no. 3, pp. 2489–2500, 2012.

[25] B. Yoon and Y. Park, "A systematic approach for identifying technology opportunities: Keyword-based morphology analysis," *Technol. Forecast. Soc. Change*, vol. 72, no. 2, pp. 145–160, 2005.

[26] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.

[27] C. Wang and D. M. Blei, "Collaborative topic modeling for recommending scientific articles," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 448–456.

[28] I. Bíró, D. Siklósi, J. Szabó, and A. A. Benczúr, "Latent Dirichlet allocation in Web spam filtering," in *Proc. 4th Int. Workshop Adversarial Inf. Retr. Web*, 2008, pp. 29–32.

[29] D. Xing and M. Girolami, "Employing latent Dirichlet allocation for fraud detection in telecommunications," *Pattern Recognit. Lett.*, vol. 28, no. 13, pp. 1727–1734, 2007.

[30] R. Krestel, P. Fankhauser, and W. Nejdl, "Latent Dirichlet allocation for tag recommendation," in *Proc. 3rd ACM Conf. Recommender Syst.*, 2009, pp. 61–68.

[31] X. Jin, Y. Zhou, and B. Mobasher, "A maximum entropy Web recommendation system: Combining collaborative and content features," in *Proc. 7th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2005, pp. 612–617.

[32] A. S. Das, M. Datar, A. Garg, and S. Rajaram, "Google news personalization: Scalable online collaborative filtering," in *Proc. 16th Int. Conf. World Wide Web*, 2007, pp. 271–280.

[33] C. H. Song, J. Yoon, N. Ko, and J. W. Han, "Mapping the technological knowledge landscape: The case of epigenetics," *Recent Patents Anti-Cancer Drug Discovery*, vol. 11, no. 4, pp. 424–433, 2016.

[34] J. Yoon, W. Seo, B.-Y. Coh, I. Song, and J.-M. Lee, "Identifying product opportunities using collaborative filtering-based patent analysis," *Comput. Ind. Eng.*, vol. 107, pp. 376–387, May 2016.

[35] R. A. Hanneman and M. Riddle, *Introduction to Social Network Methods*. Riverside, CA, USA: Univ. California Riverside, 2005.
[36] S.-S. Ko, N. Ko, D. Kim, H. Park, and J. Yoon, "Analyzing technology impact networks for R&D planning using patents: Combined application of network approaches," *Scientometrics*, vol. 101, no. 1, pp. 917–936, 2014.
[37] Y.-J. Han and Y. Park, "Patent network analysis of inter-industrial knowledge flows: The case of Korea between traditional and emerging industries," *World Patent Inf.*, vol. 28, no. 3, pp. 235–247, 2006.
[38] N. Ko, J. Yoon, and W. Seo, "Analyzing interdisciplinarity of technology fusion using knowledge flows of patents," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1955–1963, 2014.
[39] T.-S. Cho and H.-Y. Shih, "Patent citation network analysis of core and emerging technologies in Taiwan: 1997–2008," *Scientometrics*, vol. 89, pp. 795–811, Dec. 2011.
[40] B. Verspagen, "Mapping technological trajectories as patent citation networks: A study on the history of fuel cell research," *Adv. Complex Syst.*, vol. 10, no. 1, pp. 93–115, 2007.
[41] B. Wang, S. Liu, K. Ding, Z. Liu, and J. Xu, "Identifying technological topics and institution-topic distribution probability for patent competitive intelligence analysis: A case study in LTE technology," *Scientometrics*, vol. 101, no. 1, pp. 685–704, 2014.
[42] P.-C. Lee, H.-N. Su, and T.-Y. Chan, "Assessment of ontology-based knowledge network formation by vector-space model," *Scientometrics*, vol. 85, no. 3, pp. 689–703, 2010.
[43] A. W. Ulwick, *What Customers Want: Using Outcome-Driven Innovation to Create Breakthrough Products and Services*. New York, NY, USA: McGraw-Hill, 2005.
[44] C. Jeong and K. Kim, "Creating patents on the new technology using analogy-based patent mining," *Expert Syst. Appl.*, vol. 41, no. 8, pp. 3605–3614, 2014.
[45] B. Jeong, J. Yoon, and J.-M. Lee, "Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis," *Int. J. Inf. Manage.*, to be published, doi: https://doi.org/10.1016/j.ijinfomgt.2017.09.009.
[46] M. B. Albert, D. Avery, F. Narin, and P. McAllister, "Direct validation of citation counts as indicators of industrially important patents," *Res. Policy*, vol. 20, no. 3, pp. 251–259, 1991.
[47] A. Breitzman and P. Thomas, "Using patent citation analysis to target/value M&A candidates," *Res.-Technol. Manage.*, vol. 45, no. 5, pp. 28–36, 2002.
[48] F. Narin, "Technology indicators and corporate strategy," *Rev. Bus.*, vol. 14, no. 3, pp. 19–23, 1993.
[49] Y. Kajikawa and Y. Takeda, "Citation network analysis of organic LEDs," *Technol. Forecast. Soc. Change*, vol. 76, no. 8, pp. 1115–1123, 2009.
[50] T.-L. Huang and F. H. Liu, "Formation of augmented-reality interactive technology's persuasive effects from the perspective of experiential value," *Internet Res.*, vol. 24, no. 1, pp. 82–109, 2014.
[51] T.-L. Huang and S. Liao, "A model of acceptance of augmented-reality interactive technology: The moderating role of cognitive innovativeness," *Electron. Commerce Res.*, vol. 15, no. 2, pp. 269–295, 2015.
[52] S. G. Dacko, "Enabling smart retail settings via mobile augmented reality shopping apps," *Technol. Forecast. Soc. Change*, vol. 124, pp. 243–256, Nov. 2016.
[53] A. Rese, D. Baier, A. Geyer-Schulz, and S. Schreiber, "How augmented reality apps are accepted by consumers: A comparative analysis using scales and opinions," *Technol. Forecast. Soc. Change*, vol. 124, pp. 306–319, Nov. 2016.
[54] D. W. F. van Krevelen and R. Poelman, "Augmented reality: Technologies, applications, and limitations," Dept. Comput. Sci., Vrije Univ. Amsterdam, Amsterdam, The Netherlands, Tech. Rep., 2007, doi: https://doi.org/10.13140/RG.2.1.1874.7929.

**HAYOUNG CHOI** was born in Chungcheong, South Korea, in 1992. She received the B.S. and M.A. degrees in industrial engineering from Konkuk University, Seoul, South Korea, in 2016 and 2017, respectively.

She is currently a Researcher with the Department of Industrial Engineering, Konkuk University. Her research topics include technology opportunity analysis and future-oriented research and development planning.

Ms. Choi is a member of the Korean Institute of Industrial Engineers.

**SEUNGHYUN OH** was born in Seoul, South Korea, in 1993. He received the B.S. degree in industrial engineering from Konkuk University, Seoul, in 2017, where he is currently pursuing the M.S. degree.

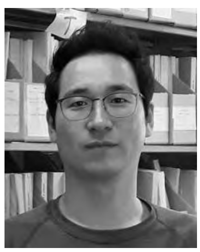
He is currently a Researcher with the Department of Industrial Engineering, Konkuk University. His research topics include computational patent mining, patent bibliometrics, and artificial intelligence-based technology analytics.

Mr. Oh is a member of the Korean Institute of Industrial Engineers.

**SUNGCHUL CHOI** was born in Busan, South Korea, in 1982. He received the B.S. degree in management from Handong Global University and the Ph.D. degree in industrial management engineering from the Pohang University of Science Technology, South Korea, in 2006 and 2012, respectively.

He was with CTO office, Samsung Advanced Institute of Technology, as a Technology Strategy Manager.He is currently an Assistant Professor with the Department of Industrial and Management Engineering, Gachon University. He has conducted research on patent analysis, technology roadmapping, and strategic planning, using a text mining approach. He was involved in applying machine learning and deep learning approaches to patent big data and has been conducting various studies related to NLP and artificial intelligence.

Dr. Choi is a member of the Korean Institute of Industrial Engineers.



**JANGHYEOK YOON** was born in Daegu, South Korea, in 1979. He received the B.S., M.S., and Ph.D. degrees in industrial engineering from the Pohang University of Science Technology, Pohang, South Korea, in 2002, 2004, and 2011, respectively.

He was an IT Consultant and Application Engineer with the LG Group from 2004 to 2007, and was an Associate Researcher with the Korea Institute of Intellectual Property, a Korean Public Research Institute, from 2011 to 2012. Since 2012, he has been an Associate Professor with the Department of Industrial Engineering, Konkuk University, Seoul, South Korea. He has authored over 50 articles, and holds 10 patents related to patent mining-based technology intelligence and opportunity discovery systems. His research interests are on technology and service intelligence, which covers the analytical and multidisciplinary methodologies for technology, service (business model) and future planning, and thus they are based on the combined application of theories and analytics in various fields, such as industrial engineering, business administration, and computer science.

Dr. Yoon is a member of the Korean Institute of Industrial Engineers. He has been an Editor of the *Journal of Intellectual Property* since 2014.