sustainability

MDPI

Article

A Hybrid Method of Analyzing Patents for Sustainable Technology Management in Humanoid Robot Industry

Jongchan Kim [1], Joonhyuck Lee [1], Gabjo Kim [2], Sangsung Park [3,*] and Dongsik Jang [1]

1 Department of Industrial Management Engineering, Korea University, Seoul 02841, Korea; ourjongchan@korea.ac.kr (J.K.); iguana751@korea.ac.kr (J.L.); jang@korea.ac.kr (D.J.)
2 Korea Intellectual Property Strategy Agency, Seoul 02841, Korea; kkjjo@korea.ac.kr
3 Graduate School of Management of Technology, Korea University, Seoul 02841, Korea
* Correspondence: hanyul@korea.ac.kr; Tel.: +82-10-8700-0469; Fax: +82-02-3290-3900

Academic Editor: Giuseppe Ioppolo
Received: 30 April 2016; Accepted: 9 May 2016; Published: 12 May 2016

Abstract: A humanoid, which refers to a robot that resembles a human body, imitates a human's intelligence, behavior, sense, and interaction in order to provide various types of services to human beings. Humanoids have been studied and developed constantly in order to improve their performance. Humanoids were previously developed for simple repetitive or hard work that required significant human power. However, intelligent service robots have been developed actively these days to provide necessary information and enjoyment; these include robots manufactured for home, entertainment, and personal use. It has become generally known that artificial intelligence humanoid technology will significantly benefit civilization. On the other hand, Successful Research and Development (R & D) on humanoids is possible only if they are developed in a proper direction in accordance with changes in markets and society. Therefore, it is necessary to analyze changes in technology markets and society for developing sustainable Management of Technology (MOT) strategies. In this study, patent data related to humanoids are analyzed by various data mining techniques, including topic modeling, cross-impact analysis, association rule mining, and social network analysis, to suggest sustainable strategies and methodologies for MOT.

Keywords: sustainable technology management; humanoid; cross-impact analysis; topic model; patents; network analysis

---

1. Introduction

A humanoid refers to a robot that resembles a human body. Humanoids have been constantly developed in order to provide various types of services as well as to perform work that is difficult for humans [1]. Humanoids were previously applied for simple repetitive or hard work that required significant human power. However, intelligent service robots for home, entertainment, and personal use have been developed actively these days to provide necessary information and enjoyment [2]. That is, they have been studied to make them resemble human behavior. Developed countries, such as the United States and Japan, have already started studying humanoids in order to dominate the robot market. A humanoid is more ideal than other types of service robots because of its unique aspects, such as its ability to communicate or interact. It is a multi-purpose worker, which makes it more efficient, as it can replace a number of special purpose robots. It would be possible to develop machines that function like a robot if a humanoid is able to handle tools or vehicles. It is better to develop a multi-functional humanoid than to develop several robots specialized for a certain purpose in terms of cost and efficiency. The application of humanoids can be making a humanoid that can handle vehicles or other types of tools.

This would imply that humanoid technology could be extended to the autonomous vehicle area. Moreover, there has been an increasing number of studies on the policies and laws for the introduction of robots [3]. It is expected that humanoids with artificial intelligence will be much more convenient for our society and its applications will appear in the near future. Thus, it is necessary to develop the strategies for sustainable management of technology (MOT) in order to predict and deal with the latent problems that may be caused by the introduction of humanoid technology. Thus, we propose sustainable humanoid MOT strategies in this paper. The development of humanoid is related to various types of technologies such as electronics, machinery, Artificial Intelligence (AI) and biotechnology. The sub-technologies of humanoid are advanced, having influence on each other. For the sustainable management of humanoid technologies, it is required to develop an effective methodology for evaluating the influence of a kind of sub-technology on the other sub-technologies as it is advanced with time [4]. Successful R & D on humanoids is possible only if it is in accordance with the changes of the market and society [5]. A cross-impact matrix is a kind of forecasting methodology that assumes future events influence on each other [6]. It is a possibility-based analyzing method that calculates the probability that when an event occurs, another event also occurs due to the influence of the event. Gordon and Hayward (1968) studied technology forecasting with cross impact analysis in order to improve the limitation of the technology forecasting based on Delphi method [7]. In the cross-impact matrix proposed in this paper, the cross-impact is given as a probability evaluated from a qualitative analysis done by a domain expert. Thus, forecasting methods were dependent on the opinions of the domain expert. However, a qualitative analysis requires too much time and cost, and it could be too subjective [8]. To improve this, quantitative forecasting methodologies need to be developed for technology forecasting. Technology forecasting can be performed with patents, known to be reliable data. A patent contains bibliographic information such as application data, applicant, international patent classification, and technological information such as title, abstract, and claims [9,10]. Currently, a quantitative patent analysis is mainly performed for obtaining basic statistics about bibliographic information and technological information is qualitatively analyzed by domain experts. A quantitative analysis, such as a statistical analysis or a data mining technique, requires a structured dataset, which is not applicable for a patent's technological information such as texts or pictures. It is currently possible to analyze unstructured data quantitatively by transforming unstructured data to structured data by using text-mining techniques. This method allows examining the technological information that has not been analyzed by domain experts [11–14]. This paper proposes a sustainable MOT strategy by applying data mining techniques to structured patent data related to humanoid technology. First, a cross-impact matrix is generated using association rule mining and a keyword analysis and the topic model is utilized to classify sub-technologies. Second, a patent analysis method using a weighted network is suggested and then the relationships between sub-technologies are visualized in order to predict the development trend of humanoid technology. In addition, an applicant analysis is performed to understand the technology market share of leading companies in the humanoid industry and the results are used to visualize the status of the technology market share by constructing a bipartite network. Finally, we suggest a sustainable MOT strategy through the two network graphs.

2. Related Work

2.1. Cross-Impact Matrix

A cross-impact matrix is a theory of predicting the future based on the fact that upcoming events influence on each other and can be used to forecast the development trend of humanoid technology [6]. A cross-impact analysis is a scenario-based analysis that produces a cross-impact matrix corresponding to a target technology. The cross impact is calculated as a conditional probability of the occurrence of an event given that its related event occurred.

Gordon (1968) proved the validity of the cross-impact matrix by carrying out a case study on a set of 28 events related to Minuteman missile arrangement [7]. Jeong and Kim (1997) proposed a method for extracting a core technology considering the cross impacts of technologies using a cross-impact matrix, the values of which were determined by experts' qualitative evaluation [15]. Choi *et al.* showed the relationship among different Information and Communications Technologies (ICT) technologies through a patent analysis that calculated the quantitative cross impacts of the technologies classified according to International Patent Classification (IPC) codes [16]. Thorleuchter *et al.* showed the relationship between the diverse defense technologies through R & D-based and patent-based cross-impact analyses [17]. However, these two papers on quantitative cross-impact analysis dealt with the criteria for patent and R & D classification that were determined by the domain experts. Therefore, this study applies cross-impact analysis on humanoid robot technology that does not include an expert's subjective technology evaluation by using a quantitative analysis technique such as a topic model.

2.2. Topic Models

A topic model is an algorithm that extracts specific topics from a group of unstructured documents [18]. It is operated by selecting a topic from clustered keywords having similar meaning within the group. There are several topics in a document group that are comprised of multiple keywords, and it leads to the extraction of specific keywords.

Zhang *et al.* (2016) studied technology forecasting by using a topic model with the United States National Science Foundation award data [19]. We extract the sub-technologies of a humanoid by applying Latent Dirichlet allocation (LDA), which is a topic model algorithm [18]. The relationship between each patent and sub-technology are identified using keywords of high frequency, term frequency-inverse document frequency (TF-IDF), and co-keyword analysis.

2.3. Association Rule Mining

Association rule mining is also known as a market basket analysis. This algorithm has been used to analyze the purchasing patterns of market customers from their baskets. There are three measurement parameters to show the relationships among different items in the basket, support, confidence, and lift. Support is defined as the proportion of item A and B at the same time. It explains the importance of an association rule between the two items [20].

$$Support\,(A \rightarrow B) = P(A \cap B) = \frac{\#\,of\ Transactions\ contain\ Item\ A\ and\ Item\ B\ at\ the\ same\ time}{\#\,of\ All\ Transactions} \quad (1)$$

Confidence is the proportion of transaction that contains item A that also contains item B; thus, it shows the possibility of purchasing item B when purchasing item A.

$$Confidence\,(A \rightarrow B) = \frac{P(A \cap B)}{P\,(A)} = \frac{Support\,(A \rightarrow B)}{\#\,of\ Transaction\ contain\ Item\ A} \quad (2)$$

Lift is defined as the observed confidence over the ratio of transactions containing item B. If a particular rule has a lift of 1, it implies that two items are independent. The two items have a positive relation when the lift is over 1, whereas they have a negative relation when the lift is below 1.

$$\mathrm{Lfit}\ \left(\mathrm{A} \rightarrow B\right) = \frac{\mathrm{P(A \cap B)}}{\mathrm{P\,(A) \times P\,(B)}} = \frac{\mathrm{Confidence\,(A \rightarrow B)}}{\mathrm{\#\ of\ Transaction\ contain\ Item\ B}} \quad (3)$$

Shih *et al.* (2008) applied the association rule mining to analyze patent and R & D trends of the Thai semiconductor industry [21]. Kim *et al.* (2011) also applied the association rule mining to analyze the relationship among different ICT technologies by measuring their cross impacts [22].

This study uses transactions as patent documents and items as sub-technologies to apply the association rule mining. The impact of each sub-technology is measured by confidence. Then, a confidence-based cross-impact matrix is constructed. However, the loss of information is inevitable because the patent analysis based on the association rule mining cannot consider the patents having just one classification. Thus, we need to compensate the loss of information by performing the keyword analysis and comparing its result with that of the cross-impact analysis.

2.4. Co-Keyword Analysis

Co-keyword analysis refers to a method for evaluating the relationships between documents according to the co-occurrence of words. Lee and Jeong (2008) suggested a strategic diagram regarding a robot technology by performing a co-keyword analysis on the national R & D data of Korea [23]. Rip and Courtial (1984) extracted the relationships between sub-technologies and analyzed the changes of the technologies over time [24].

In this study, we extract representative keywords from each sub-technology and patent document, and then derive the relationship between the patent documents and topics. Then, a cross-impact matrix is constructed with the measured impact of each sub-technology by using keyword frequency from the patent documents in each topic.

2.5. Social Network Analysis

Social network analysis is an information visualization method that represents relationships (e.g., friends, relatives, transactions, common interest, *etc.*) as nodes and edges. It has recently been applied in not only sociology but also other various fields of study. It is especially known as an efficient method to visualize the relationships among technologies.

Otte and Rousseau (2002) studied information science with social network analysis by using publication, citation, co-citation, and co-author in order to extract the relationship among information scientist, and visualized a network graph [25]. Yoon *et al.* applied natural language processing on collected patent data to extract properties and functions automatically [26]. They analyzed the technology trend of silicon-based thin film solar cells by using properties and function as nodes and suggested an invention property-function network that used co-occurrence as links.

We visualize the relationships between sub-technologies based on a weighted network, whose strength (cross impact) is shown by the thickness of edge. Especially, a bipartite network is used to show the relationships among n node groups (applicants) and other m node groups (sub-technologies). Then, we suggest a sustainable MOT strategy by using the two network graphs.

Each research dealt with the core technology extraction, the technological trends, understanding the relationships of technologies and their visualization based on the quantitative analyses. The previous studies which used the cross-impact analysis included subjective judgments because they utilized the tech-trees already determined in other reports or defined the technologies according to the opinions of domain experts. In order to improve objectiveness, we extracted and classified the specific technologies from patent documents using a quantitative approach based on Topic model. The current methods of evaluating the cross impact mainly used Association rule mining. It is inevitable that there is some loss of information when the relationships between technologies are analyzed by Association rule mining because it cannot consider the patents having just one technology classification. To overcome the limitation, the research used both Association rule mining and Co-keyword analysis to understand the relationships among technologies and tried to achieve improvement in the results by the combination. At last, the result of experiment can be understood by constructing a social network graph.

3. Hybrid Topic Model for Sustainable Technology Management

3.1. Sustainable Technology Management for Humanoids

We have seen the current level of AI in humanoids from a challenge match between Google Deepmind Alphago and Lee Sedol [27]. The challenge match has shown that AI has reached a high level of technology, which is higher than expected. Thus, it is strongly necessary to prepare sustainable MOT strategies, as a high level of AI is potentially beneficial as well as dangerous to human beings. Stanford University has observed the development of AI technology and its direction of change for hundred years, and they also studied how humanoids would influence on human life, work, and communication [28]. With the development of AI technology, humanoid technology is also being advanced rapidly. As humanoids can be applied to various fields, there are difficulties in predicting their influences on human life and determining the R & D direction. Therefore, it is necessary to develop sustainable MOT strategies in R & D of humanoid technology.

Sustainable MOT strategies enable monitoring the R & D trend of humanoid technology over time as well as suggesting flexible MOT strategies by understanding the technology market and environment analysis [29,30].

3.2. Proposed Hybrid Model

A technology consists of various sub-technologies, which have effects on each other as it is advanced. The effect needs to be understood to achieve sustainable MOT strategies. Therefore, this paper suggests a quantitative patent analysis with a hybrid model comprising text-mining, TF-IDF, topic model, cross-impact analysis, association rule mining, keyword analysis, social network analysis, and applicant analysis, as shown in Figure 1, which shows the objective and sustainable MOT strategies based on the impacts of sub-technologies.

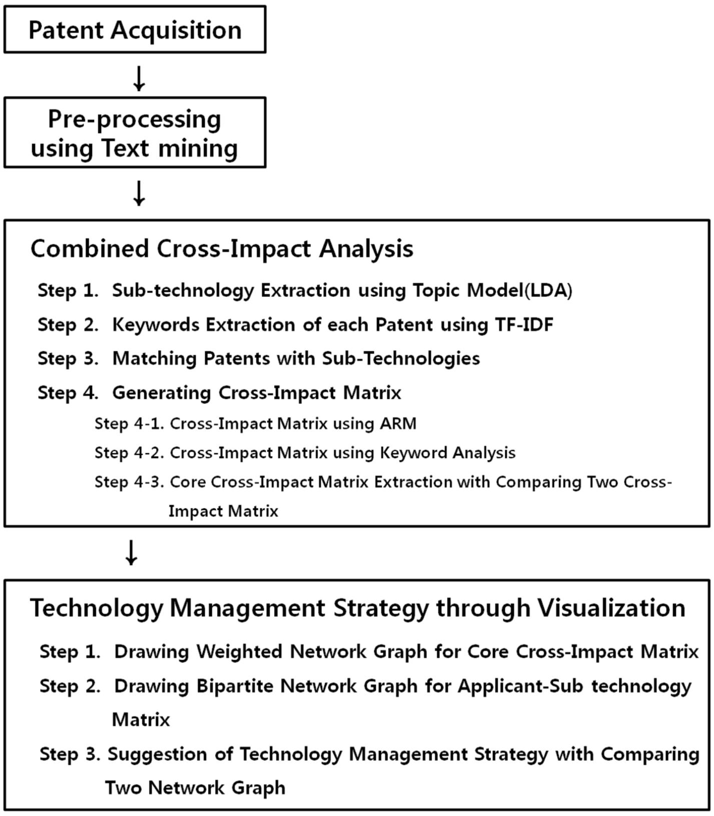

Figure 1. Procedure of proposed method.

First, patent documents are collected from the patent database. It is necessary to build the search query to collect patent data. Then, noise data are removed to select valid data. This selection step increases the accuracy of the result of analysis.

Second, a text mining technique is applied to unstructured patent data to transform the patent's title and abstract into patent-term matrix (PTM). Useless terms, such as blanks, numbers, and "be" verbs, are removed. Then, TF-IDF is used to remove the commonly occurring terms that do not have enough representativeness.

It is important to fix the number of topics in order that the topic model is applied to extract latent topics in a group of documents. We used the average silhouette width, which is mainly used in cluster analysis for extracting the optimal number of clusters [31]. Equation (4) expresses the silhouette width of each observed value.

$$S_i = \frac{D_i - M_i}{max\,\{M_i D_i\}} \tag{4}$$

where D_i refers to the average value of dissimilarity between the observed value i and the other values. M_i is the minimum distance between the value$i$ and the values in the other cluster. Thus, the largest averages silhouette width represents the optimal number of cluster.

The extracted topics from the patent group are determined as sub-technologies. Then, keywords from each topic are extracted. In order to allocate patents to the sub-technologies, it is necessary to match the keywords of the patent's high TF-IDF and sub-technologies. At this point, the patents are allowed to be allocated in more than one sub-technology. It means that a single patent can contain many technologies.

We combine two types of analysis methods in order to construct the cross-impact matrix based on the patent-sub technology matrix. First, association rule mining is applied to calculate the confidence among sub-technologies with a patent considered as a transaction and a sub-technology considered as an item. Confidence shows the possibilities of developing sub-technology B when sub-technology A has been developed. It can be defined as the cross impact between these two sub-technologies.

It is possible to evaluate the cross impact easily by applying association rule mining, but the loss of information is inevitable as it does not consider patent information with only one sub-technology. Therefore, we additionally conduct the experiment of a keyword analysis method to minimize the loss of information as well as to improve the accuracy of the experimental result. The keyword analysis method calculates the cross impact by using the ratio keywords of each sub-technology, which are included in patents.

Lastly, we use a bipartite network and visualize the relationship between high-ranked patent applicants and sub-technologies to understand the current R & D situation of companies. A sustainable MOT strategy is then finally suggested by using the two graphs.

4. Experimental Results and Analysis

4.1. Patent Acquisition

About 1000 patents regarding humanoids applied in the United States before 27 December 2015 were collected. Then, some noise data, such as patents about toys or dolls, were removed; thus, there were 901 valid raw data for the analysis, as shown in Table 1.

Table 1. Patent acquisition.

Database	Keyword	Period	Nation	Patents	Noise Data	Valid Data
Wips	Humanoid	–27 December 2015	US	1000	99	901

Wips is the name of patent database.

4.2. Pre-Processing Using Text Mining

A title and an abstract of each patent, which were unstructured, were combined as a column data. It was transformed into a PTM with 901 patents and 4415 terms by using text mining. This PTM (901 × 4415) was pre-processed and underwent TF-IDF to remove meaningless stop-words and patents. Then, a PTM with 874 patents and 1104 terms was obtained to use for further analysis.

4.3. Combined Cross-Impact Analysis

4.3.1. Sub-Technology Extraction Using Topic Model (LDA)

Extracting sub-technologies should be preceded in order to construct the cross-impact matrix. We selected 13 topics, which showed the highest average silhouette width, as shown in Table 2.

Table 2. Average value of silhouette width by a number of topics.

X = 4	**X = 5**	**X = 6**	**X = 7**	**X = 8**	**X = 9**	**X = 10**	**X = 11**	**X = 12**	**X = 13**
0.03	0.04	0.05	0.06	0.07	0.07	0.08	0.08	0.09	0.094
X = 14	**X = 15**	**X = 16**	**X = 17**	**X = 18**	**X = 19**	**X = 20**	**X = 21**	**X = 22**	**X = 23**
−0.003	0.005	0.01	0.016	0.0215	0.026	0.034	0.038	−0.052	−0.048
X = 24	**X = 25**	**X = 26**	**X = 27**	**X = 28**	**X = 29**	**X = 30**	**X = 31**	**X = 32**	**X = 33**
−0.043	−0.039	−0.35	−0.03	−0.025	−0.02	−0.015	−0.011	−0.007	−0.003
X = 34	**X = 35**	**X = 36**	**X = 37**	**X = 38**	**X = 39**	**X = 40**	**X = 41**	**X = 42**	**X = 43**
0.001	0.004	0.008	0.0127	0.0169	0.0206	0.0225	0.0267	0.0298	0.0233

Table 3 shows the keywords that belonged to 13 sub-technologies derived by topic model.

Table 3. Derived topics, keywords and technology definition.

Topic	Keywords	Technology
1	Messag, sender, voic, multimedia, surgic, microphone, document, invas, text, entity, audio, emoticon, invari, entity, instrument	Communication System
2	Platform, fall, articul, assemble, shoulder, phantom, dphm, widget, organ, tendon, heterogen, traction, depthbas, anthropomorph, centermass	Robot Joint
3	Station, section, speech, event, nois, line, convers, water, quiet, acoust, status, face, overhead, session, genom	Remote Station
4	Member, ball, guid, finger, bin, terrain, cooper, panorama, popul, cohort, bar, puzzl, counterbalance, cross, performed	Robot Finger
5	Assembl, wheel, acceler, spring, rotary, vehicle, imped, gripper, air, flexion, transmitt, articular, candi, lingual, plantar	Robotic Vehicle
6	Trunk, knowledge, touchscreen, hiup game, test, crane, neck, salienc, driver, item, loop, class, figurin, cord	Knowledge-based Robot
7	Emot, attent, tactil, easi, tag, physiology, scanner, array, illus, good, transit, stimulus, frequenc, marker, nervous	Tactile Sensor
8	Softwar, layer, hal, hardwar, finger, seat, media, grip, rnn, beam, vehicle, wheel, emphas, context, passing	Hardware Abstraction Layer
9	Pose, energy, key, exoskeleton, nonhumanoid, beat, music, slide, pipe, guid, cuff, telescop, teach, metal, descriptor	Energy Management and Sound System
10	Patient, station, batteri, gateway, breast, framework, bound, beacon plate, charg, engine, status, bridg, bright, acquisit	Patient Rounding System
11	Depth, target, zmp, gait, ann, question, shell, rout, bone, appendage, isol, grasp, bag, geography, cloud	Depth Map and Gait System
12	Anim, path, node, entity, file, face, extens, mesh, flexion, soft, creation, tissue, graph, share, brake	Path Planning System
13	Clean, fluid, subsystem, debri, avatar, assembl, world, zone, term, cash, channel, core, roller, avatars, floorclean	Cleaning Robot System

4.3.2. Keywords Extraction of Each Patent Using TF-IDF

In order to assign 874 patents to 13 sub-technologies, keywords appearing both in patents and sub-technologies were extracted and analyzed. As each keyword of different sub-technologies was extracted by the topic model, each patent's keyword was weighted according to its value of TF-IDF. Then, the words of patents whose values of TF-IDF were bigger than the third quartile, 0.17650, were selected as valid keywords.

4.3.3. Matching Patents with Sub-Technologies

Next, patents that contained keywords of each sub-technology were assigned to the corresponding sub-technologies. As a result, a PTM (776 × 13) was constructed with 776 patents, excluding 98 patents that did not belong to any sub-technology, and 13 sub-technologies. Each cell of the matrix contained 0 or 1 based on whether the patent in the row was assigned to the sub-technology in the column. Table 4 shows a part of the PTM.

Table 4. Patent-technology matrix.

Patents	Tech. 1	Tech. 2	...	Tech. 13
1	0	1	...	1
2	0	1	...	0
⋮	⋮	⋮	⋮	⋮
776	0	0	...	0

Tech.: Technology.

4.3.4. Generating Cross-Impact Matrix

Association rule mining was used to construct a cross-impact matrix. A total of 447 patents out of 776 patents, which contained more than two technologies, were set as transactions, and 13 sub-technologies were set as items. Then, the estimated confidence was used as the cross impacts among the sub-technologies to construct the cross-impact matrix.

The red-scale colors in Figure 2 represent the degree of the cross-impact as the depth of color. When the cross-impact, as a probability, is close to 1, the impact of the technology in the row on the technology in the column is big. In this research, the values of the cross-impact bigger than 0.2 were used in this research, so the color became darker as the value increases by 0.1 for better discriminability.

	Tech. 1	Tech. 2	Tech. 3	Tech. 4	Tech. 5	Tech. 6	Tech. 7	Tech. 8	Tech. 9	Tech. 10	Tech. 11	Tech. 12	Tech. 13
Tech. 1	1	0.085	0.322	0.119	0.068	0.085	0.186	0.169	0.153	0.169	0.102	0.356	0.051
Tech. 2	0.043	1	0.397	0.19	0.526	0.112	0.069	0.164	0.172	0.198	0.181	0.069	0.397
Tech. 3	0.146	0.192	1	0.077	0.1	0.146	0.108	0.092	0.115	0.438	0.077	0.2	0.146
Tech. 4	0.073	0.229	0.104	1	0.302	0.135	0.156	0.438	0.292	0.094	0.094	0.146	0.323
Tech. 5	0.032	0.492	0.105	0.234	1	0.121	0.065	0.387	0.129	0.056	0.177	0.161	0.411
Tech. 6	0.06	0.155	0.226	0.155	0.179	1	0.167	0.179	0.202	0.071	0.202	0.155	0.119
Tech. 7	0.175	0.127	0.222	0.238	0.127	0.222	1	0.206	0.143	0.127	0.063	0.222	0.143
Tech. 8	0.096	0.183	0.115	0.404	0.462	0.144	0.125	1	0.096	0.077	0.125	0.096	0.288
Tech. 9	0.103	0.23	0.172	0.322	0.184	0.195	0.103	0.115	1	0.161	0.218	0.218	0.138
Tech.10	0.112	0.258	0.64	0.101	0.079	0.067	0.09	0.09	0.157	1	0.034	0.09	0.124
Tech.11	0.072	0.253	0.12	0.108	0.265	0.205	< 0.001	0.157	0.229	0.036	1	0.229	0.06
Tech.12	0.236	0.09	0.292	0.157	0.225	0.146	0.157	0.112	0.213	0.09	0.213	1	0.09
Tech.13	0.03	0.46	0.19	0.31	0.51	0.1	0.09	0.3	0.12	0.11	0.269	0.402	1

Figure 2. Cross-impact matrix based on association rule mining. (Tech.: Technology).

The above PTM considers only the patents belonging to more than one sub-technology. As it neglected the information of 329 patents assigned to just one sub-technology, the cross-impact matrix, as shown in Table 5, was constructed considering the result of keywords analysis using all keywords.

Table 5. Result of keyword analysis.

	Tech. 1	Tech. 2	Tech. 3	Tech. 4	Tech. 5	Tech. 6	Tech. 7	Tech. 8	Tech. 9	Tech. 10	Tech. 11	Tech. 12	Tech. 13
Tech. 1	6.605	0.099	1.025	0.321	0.086	0.272	0.383	0.148	0.605	0.506	0.198	2.012	0.049
Tech. 2	0.17	4.106	0.716	0.596	2.305	0.262	0.135	0.511	0.255	0.645	0.525	0.497	2.234
Tech. 3	0.773	0.633	3.987	0.14	0.407	0.24	0.207	0.2	0.44	1.807	0.167	1.12	0.46
Tech. 4	0.288	0.586	0.297	4.496	0.784	0.279	0.405	1.27	1.243	0.261	0.261	0.469	0.991
Tech. 5	0.095	2.157	0.401	0.796	4.122	0.306	0.15	2.32	0.374	0.095	0.605	0.469	2.395
Tech. 6	0.16	0.443	0.632	0.142	0.415	3.226	0.16	0.349	0.415	0.104	0.519	0.953	0.943
Tech. 7	0.422	0.177	0.52	0.412	0.226	0.441	3.422	0.43	0.108	0.284	0.019	0.382	0.284
Tech. 8	0.426	0.754	0.287	1.787	1.689	0.32	0.344	4.91	0.221	0.131	0.279	0.303	1.254
Tech. 9	0.4	0.52	0.38	1.4	0.47	0.72	0.24	0.34	4.22	0.31	0.65	0.45	0.93
Tech. 10	0.784	0.469	2.378	0.162	0.171	0.207	0.144	0.216	0.342	4.667	0.108	0.27	0.441
Tech. 11	0.085	0.386	0.242	0.34	0.588	0.248	0.026	0.288	0.758	0.052	4.392	0.412	0.118
Tech. 12	2.252	0.162	1.036	0.487	0.784	0.487	0.423	0.207	1.604	0.478	0.901	5.18	0.225
Tech. 13	0.131	2.131	0.516	1.254	2.189	0.197	0.238	0.885	0.197	0.73	0.189	0.197	6.762

Tech.: Technology.

Each cell P_{ij} of Table 5 denotes the average value that was calculated by the appearance frequency of the keyword of sub-technology j in patents assigned to sub-technology i divided by the number of patents. The value was transformed into the cross impact by using Equation (5), as shown in Figure 3.

$$\mathrm{P}\,(j|i) = \frac{P(i \cap j)}{P\,(j)} \rightarrow P_{ij} = \frac{P_{ij}}{P_{jj}} \tag{5}$$

The above two cross-impact matrices were combined, and then the relationship between the sub-technologies was extracted according to whether the cross impact was higher than 0.2. The cross impact lower than 0.2 was transformed into 0. The cross impact higher than 0.2 was recalculated as the average value of the two matrices. As a result, the core cross-impact matrix, which represented the combined matrix, was constructed to present the relationships only among significant sub-technologies, as shown in Table 6.

	Tech. 1	Tech. 2	Tech. 3	Tech. 4	Tech. 5	Tech. 6	Tech. 7	Tech. 8	Tech. 9	Tech. 10	Tech. 11	Tech. 12	Tech. 13
Tech. 1	1	0.024	0.257	0.071	0.021	0.084	0.112	0.03	0.143	0.108	0.045	0.388	0.007
Tech. 2	0.026	1	0.18	0.133	0.559	0.081	0.039	0.104	0.06	0.138	0.119	0.096	0.33
Tech. 3	0.117	0.154	1	0.031	0.099	0.074	0.06	0.041	0.104	0.387	0.038	0.216	0.068
Tech. 4	0.044	0.143	0.075	1	0.19	0.087	0.118	0.259	0.295	0.056	0.059	0.09	0.147
Tech. 5	0.014	0.525	0.101	0.177	1	0.095	0.044	0.472	0.089	0.02	0.138	0.091	0.354
Tech. 6	0.024	0.108	0.159	0.031	0.101	1	0.047	0.071	0.098	0.022	0.118	0.184	0.14
Tech. 7	0.064	0.043	0.13	0.092	0.055	0.137	1	0.088	0.026	0.061	0.004	0.074	0.042
Tech. 8	0.065	0.184	0.072	0.397	0.41	0.099	0.101	1	0.052	0.028	0.063	0.059	0.185
Tech. 9	0.061	0.127	0.095	0.311	0.114	0.223	0.07	0.069	1	0.066	0.148	0.087	0.138
Tech.10	0.119	0.114	0.597	0.036	0.042	0.064	0.042	0.044	0.081	1	0.025	0.052	0.065
Tech.11	0.013	0.094	0.061	0.076	0.143	0.077	0.008	0.059	0.18	0.011	1	0.079	0.017
Tech.12	0.341	0.039	0.26	0.108	0.19	0.151	0.124	0.042	0.38	0.102	0.205	1	0.033
Tech.13	0.02	0.519	0.13	0.279	0.531	0.061	0.069	0.18	0.047	0.156	0.043	0.038	1

Figure 3. Cross-impact matrix based on keyword analysis. (Tech.: Technology).

The value of thresholds 0.2 was arbitrarily set to consider the cross-impact only when the values of it is bigger than 0.2. As the thresholds decreases, we can consider more relationships between technologies, but the complexity also increases. Conversely, the number of relationships to consider decreases as the value of thresholds increases, so the relationships can be more easily understood. The problem of setting the threshold value is not that of right or wrong, so the analysts or the managers who try to get insights for sustainable MOT through the methodology proposed in this paper can adjust the value according to their purpose of analysis.

Table 6. Core cross-impact matrix.

	Tech. 1	Tech. 2	Tech. 3	Tech. 4	Tech. 5	Tech. 6	Tech. 7	Tech. 8	Tech. 9	Tech. 10	Tech. 11	Tech. 12	Tech. 13
Tech. 1	1	0	0.29	0	0	0	0	0	0	0	0	0.372	0
Tech. 2	0	1	0.289	0	0.543	0	0	0	0	0	0	0	0.364
Tech. 3	0	0	1	0	0	0	0	0	0	0.413	0	0.208	0
Tech. 4	0	0	0	1	0.246	0	0	0.349	0.294	0	0	0	0.235
Tech. 5	0	0.509	0	0.206	1	0	0	0.43	0	0	0	0	0.383
Tech. 6	0	0	0	0	0	1	0	0	0	0	0	0	0
Tech. 7	0	0	0	0	0	0	1	0	0	0	0	0	0
Tech. 8	0	0	0	0.401	0.436	0	0	1	0	0	0	0	0.237
Tech. 9	0	0	0	0.317	0	0.209	0	0	1	0	0	0	0
Tech. 10	0	0	0.619	0	0	0	0	0	0	1	0	0	0
Tech. 11	0	0	0	0	0.204	0	0	0	0.205	0	1	0	0
Tech. 12	0.289	0	0.276	0	0.208	0	0	0	0.297	0	0.209	1	0
Tech. 13	0	0.49	0	0.295	0.521	0	0	0.24	0	0	0	0.22	1

Tech.: Technology.

4.4. Combined Cross-Impact Analysis

4.4.1. Drawing Weighted Network Graph for Core Cross-Impact Matrix

The relationships between sub-technologies extracted using the core cross-impact matrix were visualized by a weighted network graph that represents the relationship of sub-technologies as edge's direction and thickness.

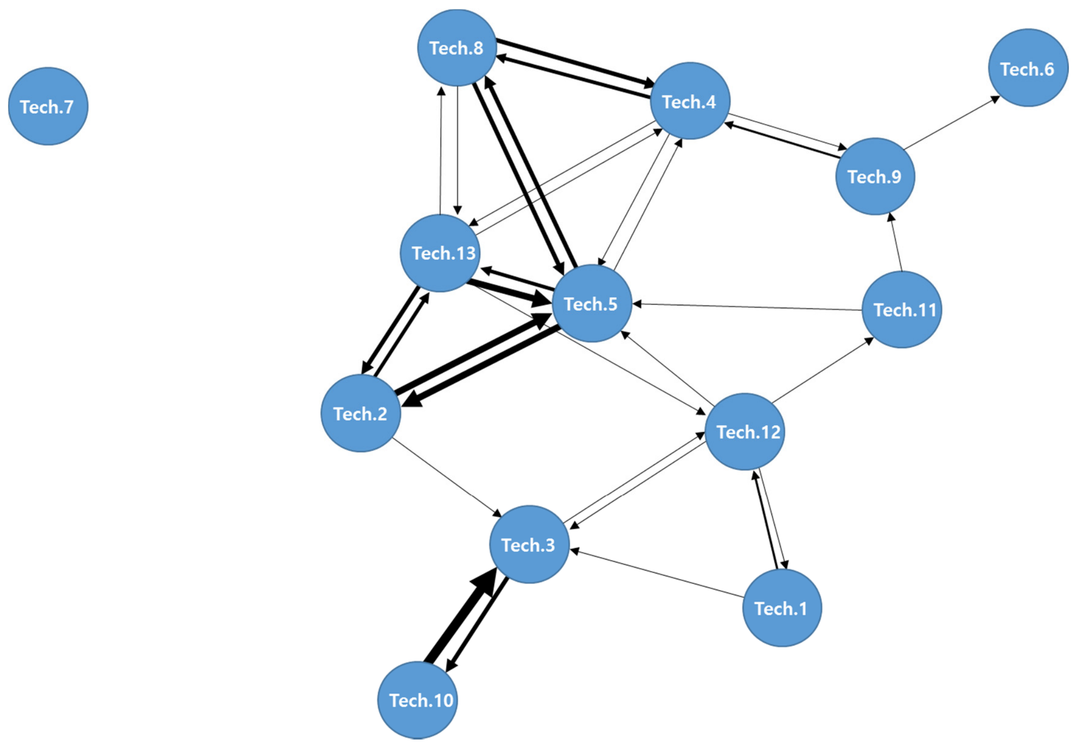

Figure 4. Weighted network graph for core Cross-impact matrix. (Tech.: Technology).

According to Figure 4, Tech. 7 and Tech. 6 are considered to be independently developed while they do not have effects on other technologies. Thus, when the developments of those two technologies are considered, other technologies do not need to be considered. Conversely, Tech. 10 and Tech. 3

have the biggest effects on each other. Seeing the fact that the development of Patient rounding system technology has a big effect on the development of Remote station technology, the latter one seems to be the core technology for the former one. Thus, the increase of the demand of the patient rounding system technology is expected to lead to the increase of the demand of the remote system technology. It can be seen that Robotic vehicle technology located in the center of the graph is linked to the biggest number of technologies. This means that the technology is the complex related to various other technologies. Therefore, the current status of various other technologies should be considered when the R & D of the robotic vehicle technology is processed. In addition, it can be found out that the technology of autonomous driving using an AI robot is a critical issue in the current humanoid robot technology. There is also a remarkable point about path planning system (Tech. 12). Seeing that it is located in the center around robotic vehicle, depth map and gait system (Tech. 11), remote system and communication system (Tech. 1), it seems to be the core technology of the humanoid mobility system. Therefore, when the humanoid mobility system is under development, the development of the path planning system and its surrounding technologies such as robotic vehicle, depth map and gait system, remote system and communication system should be considered together.

4.4.2. Drawing Bipartite Network Graph for Applicant-Sub-Technology Matrix

It is possible to understand the cross impacts and the relationships between technologies from the weighted network graph. It is also necessary to analyze the R & D companies to understand the current R & D situation of each technology, as shown in Table 7.

Table 7. Leading companies extracted.

Ranking	Company	Patents
1	Honda Research Institute Europe GmbH	100
2	iRobot Corporation	77
3	Sony Corporation	48
4	Microsoft Corporation	41
5	InTouch Technologies, Inc.	39
6	Samsung Electronics	26
7	GM Global Technology Operations, Inc.	20
8	Disney Enterprises, INC., A Delaware Corporation	19
9	Primesense Ltd.	17
10	Massachusetts Institute of Technology	16

Table 8 shows the company–technology matrix to find out the technological area where each company focuses its R & D resource. Each cell presents the number of patents owned by each company.

Table 8. Company–technology matrix.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13
Honda	17	23	30	15	13	11	10	17	15	13	36	10	0
iRobot	0	14	19	10	34	0	0	29	0	25	0	0	117
Sony	0	0	0	0	0	18	0	17	0	0	16	0	0
Microsoft	0	0	0	0	0	14	0	0	0	0	32	11	0
Intouch	0	14	38	0	0	0	0	0	0	44	0	0	0
Samsung	0	0	0	25	0	0	0	11	0	0	0	0	0
GM	0	17	0	0	21	0	0	0	0	0	0	0	13
Disney	0	0	0	0	0	0	0	0	21	0	0	0	0
Primesense	0	0	0	0	0	0	0	0	0	0	17	0	0
MIT	0	0	0	0	25	0	0	0	0	0	0	10	0

T: Technology.

The bipartite network graph was drawn based on the constructed cross-impact matrix in order to visualize and understand these companies, as shown in Figure 5.

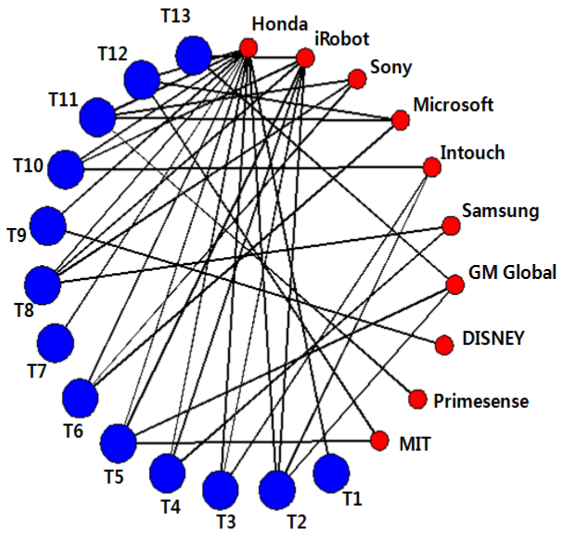

Figure 5. Bipartite network graph for company–technology matrix. (T: Technology).

It was possible to understand that Honda and iRobot actively developed various sub-technologies. Honda had many patents on almost every technological area except tech. 13(Cleaning robot system), while iRobot mostly concentrated on developing tech. 13. Other companies had patents on 1 to 3 sub-technologies each.

5. Conclusions

In this study, humanoid patents were collected to extract sub-technologies based on a topic model, and the patents were assigned to sub-technologies by TF-IDF. Then, a core cross-impact matrix was constructed using association rule mining and a keyword analysis. A weighted network was then constructed to visualize the cross impacts and the relationships between sub-technologies from the core cross-impact matrix. Additionally, a company–technology analysis was conducted and visualized using a bipartite network to understand the current R & D situation of each company. The patent analysis method presented in this study is useful for developing a sustainable MOT strategy, as it is possible to understand specific technologies. A technology consists of various types of sub-technologies that positively influence on each other. Therefore, it is strongly necessary to extract the relationships between these sub-technologies. Moreover, this paper provides additional insights into MOT by presenting a company's R & D situation. In a further study, a simulation using the cross-impact matrix will be conducted to predict promising technologies over time.

Acknowledgments: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2015R1D1A1A01059742) and by the BK 21 Plus (Big Data in Manufacturing and Logistics Systems, Korea University).

Author Contributions: Kim designed this research and collected the data set for the experiment. Lee, Kim and Jang analyzed the data to show the validity of this study. Park wrote the paper and performed the entire research steps. In addition, all authors have cooperated with each other in revising the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ting, C.H.; Yeo, W.H.; King, Y.J.; Chuah, Y.D.; Lee, J.V.; Khaw, W.B. Humanoid robot: A review of the architecture, applications and future trend. *Res. J. Appl. Sci. Eng. Technol.* **2014**, *7*, 1364–1369.

2. You, B.J.; Oh, Y.H.; Choi, Y.J. Survey on humanoid researches. *J. Korean Soc. Precision Eng.* **2004**, *21*, 15–21. (In Korean)
3. Lin, P.; Abney, K.; Bekey, G.A. *Robot Ethics: The Ethical and Social Implications of Robotics*; The MIT Press: Massachusetts, USA, 2012.
4. Lee, S.; Kang, S.; Park, Y.S.; Park, Y. Technology roadmapping for R & D planning: The case of the Korean parts and material industry. *Technovation* **2007**, *27*, 433–445.
5. Arora, A.; Gambardella, A. The market for technology. *Handb. Econ. Innov.* **2011**, *1*, 641–678.
6. Gordon, T.J. Cross-impact matrices: An illustration of their use for policy analysis. *Futures* **1969**, *1*, 527–531. [CrossRef]
7. Gordon, T.J.; Hayward, H. Initial experiments with the cross impact matrix method of forecasting. *Futures* **1968**, *1*, 100–116. [CrossRef]
8. Mishra, S.; Deshmukh, S.G.; Vrat, P. Matching of technological forecasting technique to a technology. *Technol. Forecast. Soc. Change* **2002**, *69*, 1–27. [CrossRef]
9. KIPRIS. Korea Intellectual Property Rights Information Service. Available online: http://www.kipris.or.kr (accessed on 10 May 2016).
10. Korea Intellectual Property Office. *Patent and Information Analysis (for Researchers)*; Gyungsung Book: Seoul, Korea, 2009. (In Korean)
11. Tseng, Y.H.; Lin, C.J.; Lin, Y.I. Text mining technique for patent analysis. *Inf. Processing Manage.* **2007**, *43*, 1216–1247. [CrossRef]
12. Fattori, M.; Pedrazzi, G.; Turra, R. Text mining applied to patent mapping: A practical business case. *World Patent Inf.* **2003**, *25*, 335–342. [CrossRef]
13. Yoon, B.; Park, Y. A systematic approach for identifying technology opportunities: Keyword-based morphology analysis. *Technol. Forecast. Soc. Change* **2005**, *72*, 145–160. [CrossRef]
14. Lee, S.; Yoon, B.; Park, Y. An approach to discovering new technology opportunities: Keyword-based patent map approach. *Technovation* **2009**, *29*, 481–497. [CrossRef]
15. Jeong, G.H.; Kim, S.H. A qualitative cross-impact approach to find the key technology. *Technol. Forecast. Soc. Change* **1997**, *55*, 203–214. [CrossRef]
16. Choi, C.; Kim, S.; Park, Y. A patent-based cross impact analysis for quantitative estimation of technological impact: The case of information and communication technology. *Technol. Forecast. Soc. Change* **2007**, *74*, 1296–1314. [CrossRef]
17. Thorleuchter, D.; Van den Poel, D.; Prinzie, A. A compared R & D-based and patent-based cross impact analysis for identifying relationships between technologies. *Technol. Forecast. Soc. Change* **2010**, *77*, 1037–1050.
18. Blei, D.M. Probabilistic topic models. *Commun. ACM* **2012**, *55*, 77–84. [CrossRef]
19. Zhang, Y.; Zhang, G.; Chen, H.; Porter, A.L.; Zhu, D.; Lu, J. Topic analysis and forecasting for science, technology and innovation: Methodology with a case study focusing on big data research. *Technol. Forecast. Soc. Change* **2016**, *105*, 179–191. [CrossRef]
20. Kotsiantis, S.; Kanellopoulos, D. Association rules mining: A recent overview. *Int. Trans. Comput. Sci. Eng.* **2006**, *32*, 71–82.
21. Shih, M.J.; Liu, D.R.; Hsu, M.L. Mining change in patent trends for competitive intelligence. In *Advances in Knowledge Discovery and Data Mining*; Washio, T., Suzuki, E., Ting, K.M., Inokuchi, A., Eds.; Springer: Berlin, Germany, 2000; Volume 5012, pp. 999–1005.
22. Kim, C.; Lee, H.; Seol, H.; Lee, C. Identifying core technologies based on technological cross-impacts: An association rule mining (ARM) and analytic network process (ANP) approach. *Expert Syst. Appl.* **2011**, *38*, 12559–12564. [CrossRef]
23. Lee, B.; Jeong, Y.I. Mapping Korea's national R & D domain of robot technology by using the co-word analysis. *Scientometrics* **2008**, *77*, 3–19.
24. Rip, A.; Courtial, J. Co-word maps of biotechnology: An example of cognitive scientometrics. *Scientometrics* **1984**, *6*, 381–400. [CrossRef]
25. Otte, E.; Rousseau, R. Social network analysis: A powerful strategy, also for the information sciences. *J. Inf. Sci.* **2002**, *28*, 441–453. [CrossRef]
26. Yoon, J.; Choi, S.; Kim, K. Invention property-function network analysis of patents: A case of silicon-based thin film solar cells. *Scientometrics* **2010**, *86*, 687–703. [CrossRef]

27. Silver, D.; Huang, A.; Maddison, C.J.; Guez, A.; Sifre, L.; van den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; *et al.* Mastering the game of Go with deep neural networks and tree search. *Nature* **2016**, *529*, 484–489. [CrossRef] [PubMed]
28. One Hundred Year Study on Artificial Intelligence (AI100). Available online: http://ai100.stanford.edu/ (accessed on 10 May 2016).
29. Choi, J.; Jang, D.; Jun, S.; Park, S. A predictive model of technology transfer using patent analysis. *Sustainability* **2015**, *7*, 16175–16195. [CrossRef]
30. Kim, S.; Jang, D.; Jun, S.; Park, S. A novel forecasting methodology for sustainable management of defense technology. *Sustainability* **2015**, *7*, 16720–16736. [CrossRef]
31. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [CrossRef]