February 2017;1(2):63-88
http://dx.doi.org/10.20522/APJBR.2017.1.2.63

Cross-national Analysis of Robot Research Using Non-Structured Text Analytics for R&D Policy

Jeong Hun Kim
School of Management, Kyung Hee University

Han Sol Seo
School of Management, Kyung Hee University

Jae Woong Lee
School of Management, Kyung Hee University

Jung Won Lee
School of Management, Kyung Hee University

Oh Byung Kwon*
School of Management, Kyung Hee University

Abstract

With the advent of new frontiers in robotics, the spectrum of robot research area has widened in many fields and applications. Other than conventional robot research, many technologies such as smart devices, drones, healthcare robots, and soft robots are emerging as promising applications. Due to the research complexity of this topic, this research requires international collaboration and should be fertilized by R&D policies. This paper aims to propose a method to perform a cross-national analysis of robot research with unstructured data such as papers in the proceedings of an international conference. Text analytics are applied to extract research issues and applications in an automatic manner.

Keywords Cross-national analysis; Robotics, Text analytics; TF-IDF; R&D policy

*Corresponding author: obkwon@khu.ac.kr

1. Introduction

As interest in robots has recently been increasing, the volume of the robot market is rapidly growing. According to World Robotics 2015, China, Japan, the United States, Korea, and Germany control 70% of the world robot market. China holds the largest share in the market with 25% of the world market volume. Like the market volume, research on robots is rapidly progressing and research issues are rapidly changing. Such changes in research issues could be found in the changes of topics and content of papers from conferences on robots. Papers written by researchers are unstructured data that contain a variety of academic information including the topic of the research, trends in precedent research, and the logic of the researcher. Hence, research papers could be utilized as useful material to observe the trends of the field.

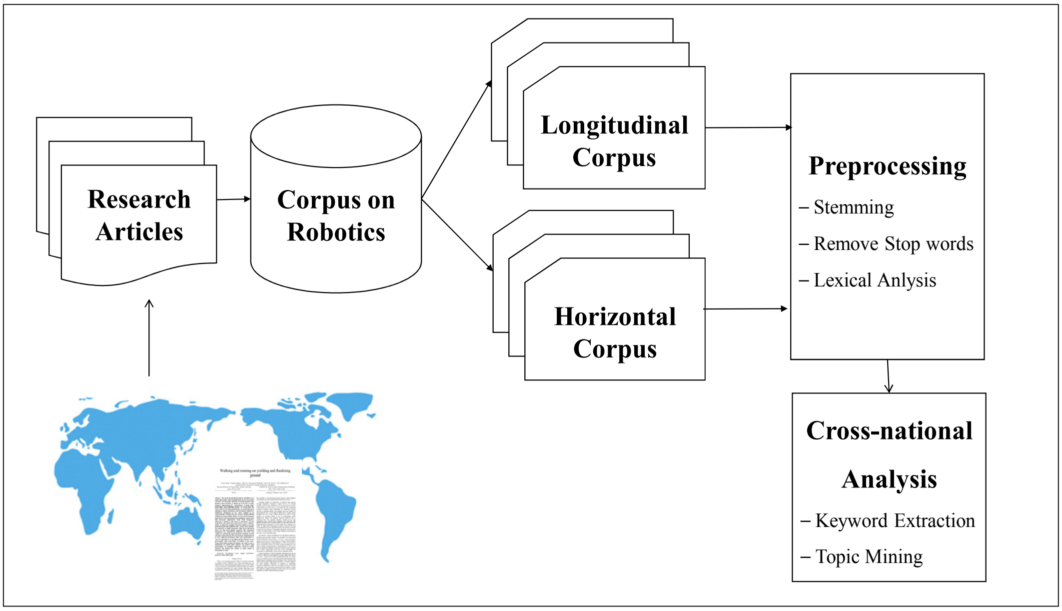
So far, research trends in robotics were mainly focused on specific domains such as research on disabled children (Kang et al. 2013; Kim and Sin, 2014), robot education (Kim, 2012), control robots (Lee et al, 2013; Lee and Jeong, 2015), and life support robots (Park et al, 2013). However, general trend analyses on overall robotics are rarely conducted.

Also, for the analysis methodology designed to understand trends, a literature study and analysis of contents by researchers (Kim, 2012) and a statistical analysis of a bibliography analysis and a social network analysis (Lee and Jeong, 2015) are used. The abstract of the research paper summarizes the overall research results in a fundamental and concise manner so that readers can quickly find the overall flow and topic of the paper (Griffiths and Steyvers, 2004). Clearly, research papers could be utilized as good material to explore the research trends in a field. With text mining, it is possible to extract information regarding various subjects. The text mining technique is a method that can automatically extract topics or main issues from massive unstructured data including papers or research reports and then visualize, categorize, or make predictions regarding this data. Text mining is an effective method in trend analysis in an overall research field, similar to this study. However, there is no study that uses the text mining method in research trend analyses on robotics or introduces robotics R&D policy based on meta-studies.

This study aims to suggest a direction for policy through a comparison of robotics research topics based on international conference papers, especially those based in Asia. For this, the study applied text mining technology to not only rely on the keywords of each paper but to extract keywords from abstracts of papers. Because text mining could find interesting but closet facts from unstructured data (Griffiths and Steyvers, 2004), it is utilized in various fields including bioinformatics (Song and Kim, 2012). Also, based on the extracted keywords, we searched for trends in core detail research fields by region and explored the relevant changes and dimensions.

Through this, the study would be able to find the research trends in robotics in each country, information which has the potential to be utilized in various studies as the basis for future studies.

Figure 1 Overall process



2. Method

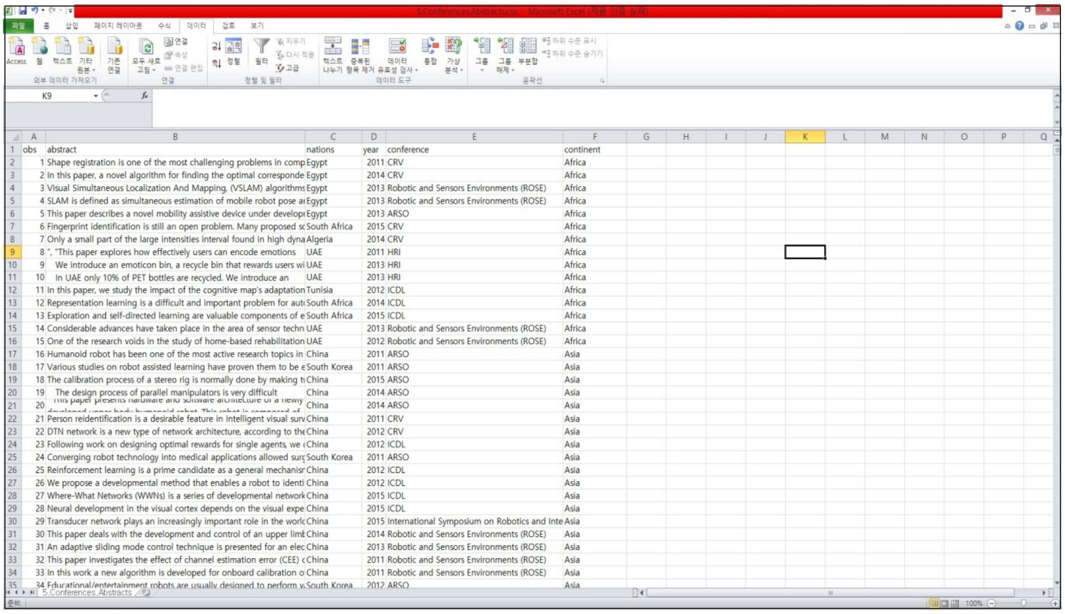
2.1 Overall process

Figure 1 is the basis for conducting a multinational analysis through collecting, arranging, and analyzing robotics research. First, we searched for robotics conference papers around the world through the Internet. For that, we utilized robotics conference information websites and Google Scholar. Then we collected the abstracts of those papers to construct a robotics corpus. Next, in accordance with the needs of the analysis, this corpus was divided into a longitudinal corpus that was made by categorizing the robotics corpus by time and a horizontal corpus that was made by categorizing the robotics corpus by country or region. Next, for the refined meta-research, preprocessing on each corpus is conducted. This includes acquiring stem words by analyzing morphemes (stemming), eliminating meaningless words (stop words) for analysis, and analyzing morphemes. After that, a cross-national analysis is conducted by region and time. For this, keywords related to robotics are extracted and topic modeling is conducted based on that.

2.2 Data collection for corpus on robotics

In order to conduct a meta-study in the literature study method suggested above, papers presented in five representative international robotics conferences such as ARSO (Advanced Robotics and its Social Impacts), CRV (Computer and Robot Vision), HRI (Human-Robot Interaction), ICDL (ICDL-EpiRob), RO-MAN (Robot and Human Interactive Communication) were selected as the analysis objects. Using a website (http://ieeexplore.ieee.org/) where the users could access journals, conference proceedings,

Figure 2 An example of the corpus on robotics



	A	B	C	D	E	F
1	obs	abstract	nations	year	conference	continent
2	1	Shape registration is one of the most challenging problems in comp	Egypt	2011	CRV	Africa
3	2	In this paper, a novel algorithm for finding the optimal corresponde	Egypt	2014	CRV	Africa
4	3	Visual Simultaneous Localization And Mapping, (VSLAM) algorithms	Egypt	2013	Robotic and Sensors Environments (ROSE)	Africa
5	4	SLAM is defined as simultaneous estimation of mobile robot pose ar	Egypt	2013	Robotic and Sensors Environments (ROSE)	Africa
6	5	This paper describes a novel mobility assistive device under develop	Egypt	2013	ARSO	Africa
7	6	Fingerprint identification is still an open problem. Many proposed sc	South Africa	2015	CRV	Africa
8	7	Only a small part of the large intensities interval found in high dyna	Algeria	2014	CRV	Africa
9	8	", "This paper explores how effectively users can encode emotions	UAE	2011	HRI	Africa
10	9	We introduce an emoticon bin, a recycle bin that rewards users wi	UAE	2013	HRI	Africa
11	10	In UAE only 10% of PET bottles are recycled. We introduce an	UAE	2013	HRI	Africa
12	11	In this paper, we study the impact of the cognitive map's adaptation	Tunisia	2012	ICDL	Africa
13	12	Representation learning is a difficult and important problem for aut	South Africa	2014	ICDL	Africa
14	13	Exploration and self-directed learning are valuable components of e	South Africa	2015	ICDL	Africa
15	14	Considerable advances have taken place in the area of sensor techn	UAE	2013	Robotic and Sensors Environments (ROSE)	Africa
16	15	One of the research voids in the study of home-based rehabilitation	UAE	2012	Robotic and Sensors Environments (ROSE)	Africa
17	16	Humanoid robot has been one of the most active research topics in	China	2011	ARSO	Asia
18	17	Various studies on robot assisted learning have proven them to be e	South Korea	2011	ARSO	Asia
19	18	The calibration process of a stereo rig is normally done by making t	China	2015	ARSO	Asia
20	19	The design process of parallel manipulators is very difficult	China	2014	ARSO	Asia
21	20	This paper presents hardware and software architecture of a newly developed upper body humanoid robot. This robot is composed of	China	2014	ARSO	Asia
22	21	Person reidentification is a desirable feature in intelligent visual surv	China	2011	CRV	Asia
23	22	DTN network is a new type of network architecture, according to the	China	2012	CRV	Asia
24	23	Following work on designing optimal rewards for single agents, we	China	2012	ICDL	Asia
25	24	Converging robot technology into medical applications allowed surg	South Korea	2011	ARSO	Asia
26	25	Reinforcement learning is a prime candidate as a general mechanisr	China	2012	ICDL	Asia
27	26	We propose a developmental method that enables a robot to identi	China	2012	ICDL	Asia
28	27	Where-What Networks (WWNs) is a series of developmental network	China	2015	ICDL	Asia
29	28	Neural development in the visual cortex depends on the visual expe	China	2015	ICDL	Asia
30	29	Transducer network plays an increasingly important role in the worl	China	2015	International Symposium on Robotics and Inte	Asia
31	30	This paper deals with the development and control of an upper limb	China	2014	Robotic and Sensors Environments (ROSE)	Asia
32	31	An adaptive sliding mode control technique is presented for an elec	China	2013	Robotic and Sensors Environments (ROSE)	Asia
33	32	This paper investigates the effect of channel estimation error (CEE) c	China	2011	Robotic and Sensors Environments (ROSE)	Asia
34	33	In this work a new algorithm is developed for onboard calibration o	China	2011	Robotic and Sensors Environments (ROSE)	Asia
35	34	Educational/entertainment robots are usually designed to perform v	South Korea	2012	ARSO	Asia

standards issued by IEEE (Institute of Electrical and Electronics Engineers), papers related to robots during the 5-year period from 2011 to 2015 were collected. PDF files were saved in the format of 'Year_Nation_Number.pdf' and were stored separately by conferences. Files of collected conference papers were in .pdf format. After downloading a paper, we used R program to extract the abstracts only to save them in DB. The information of the country of the research paper was input manually based on the institute to which the lead author belonged.

2.3 Preprocessing

PDF files were saved as 'Year_Nation_Number.pdf' and were stored by conference. In order to extract the abstracts only from saved PDF files, we used pdftotext.exe, a PDF conversion program provided by 'Foolabs' and R, an open source statistics program provided by 'R-Project.' After converting all the saved PDF files in the conference folders into text files, we checked specific patterns where the abstracts were included (e.g., between words 'Abstract' and 'Keywords') and used that pattern to extract only the abstracts from papers. Also, for the analysis by year or continent, we extracted the year and country information of a PDF file from the 'Year_Nation_Number.pdf' format, which was used as the file name, and then organized the information into a dataset, and the continent information was stored by creating variables based on the extracted countries. Figure 2 is an example of the dataset. The formed dataset was divided by year and continent for the analysis, resulting in the corpus.

Table 1 Robot research papers by conference

Conference/Year	2011	2012	2013	2014	2015	Total
ARSO	21	21	42	25	29	138
CRV	53	68	52	52	44	269
HRI	104	145	130	155	149	683
ICDL	67	101	53	85	61	367
RO-MAN	84	173	168	184	137	746
Total	329	508	445	501	420	2,203

Table 2 Robot research papers by continent

Continent/Year		2011	2012	2013	2014	2015	total
Asia	Korea	19	17	65	21	18	140
	China	3	4	2	4	4	17
	Japan	53	103	118	69	87	430
	Others	8	21	14	16	8	67
EU		114	172	125	240	140	791
North America		125	179	105	133	151	693
Africa		2	1	4	3	2	12
Oceania		5	7	7	10	6	35
South America		-	4	5	5	4	18
Total		329	508	445	501	420	2,203

Preprocessing process was conducted by 'tm,' the text mining package of 'R.' After eliminating punctuation marks and numbers, stemming was conducted and then stop words were removed. For stop words, after removing the overlapping 889 words out of the 572 words from the SMART information retrieval system, 429 Onix Text Retrieval Toolkit stopwords, 667 Ranksnl Long stopwords, and 635 Webconfs stopwords, we used 1,413 stopwords in total.

3. Result

3.1 Simple statistics

Papers collected in accordance with the methodology of this study are in Table 1 below. Papers from 2011 to 2015 were collected. Also, Table 2 is the result of categorizing papers based on the country based of the lead author. Many papers were written in the EU, North America, and Japan, and for a large number of papers from China, the lead author was frm the US or the EU.

Looking at Table 2, we can find that the number of papers is not increasingly consistently or stable but shows a large deviation by year. If we take a look at Appendix A, we can find that it is because neighboring countries present more papers based on the venue of the conference.

As papers from Asia except Korea, China, and Japan were too few, we summarized them and categorized them into Others. Also, papers from Oceania, Africa, or South America were few.

3.2 Keyword comparison by continent

Overall Core Research Field Analysis

For the keyword comparison by continent or period, there is a need for extraction of keywords in robotics research. For that, TF-IDF weights of 8,584 keywords extracted by the preprocessing of 2,203 documents in collected corpus were calculated. Here, TF-IDF weight followed the method of calculating in the order of forming a document-term matrix, summing weight value of each keyword, and organizing, as shown in Table 3. The top 100 words with the highest weights were recognized as robotics research-related keywords. Table 4 shows the results of calculating TF-IDF out of all the robotics research papers without the consideration of the collected year, and Figure 6 shows the results of calculating TF-IDF weight by year from all robotics research papers and showing the top 10 words in a graph.

Although years 2013 and 2015 show a declining flow in the number of research papers through fewer research papers compared to other years, the top 10 words still remain highly ranked. Considering the words in the top ranks such as 'interact,' 'learn,' 'social,' 'children,' 'motion,' 'gestur,' 'imag,' 'visual,' 'featur,' and 'interfac,' it seems that research on human-robot interaction (HRI) was actively conducted from 2011 to 2015. And as fields using robots, it seems that education and children have been actively discussed. Moreover, we can see that research on the circumstantial judgment for facilitating the interaction between robot and human has been actively conducted.

Figure 3, 4 are graphs that show the words that have shown the greatest increase or decrease in 2015 compared to 2011. First, in Figure 3, considering that words such as 'children,' 'program,' 'lead,' and 'relationship' appear, it is possible to see that HRI-related research are was receiving constant attention compared to 2011. Also, considering that 'autonomi' did not appear in 2011 but did in 2015, it could be interpreted as a likely rising research issue. Next, in Figure 4, we explored the words with the largest decrease in 2015 compared to 2011. We could consider words such as 'touch,' 'request,' 'dialog,' and 'represent' to be closely related to the interactions between robot and human. A decrease in the words related to the interaction between robot and human drew our attention. However, considering that words related to the HRI appear in Figure 3, we concluded that researchers were less interested in interaction or less research was being done because the quantity of research on the topic was already substantial rather than concluding that the amount of research related to interaction had seen a decrease.

Next, we analyzed the trends of the extracted top 10 words by continent. In general, 10 words were most frequently mentioned in research papers in each continent. However, in China, unlike overall research trends, not only were the TF-IDF values of the top 10 words small but the trends in the words were also uneven. It seems that this is because the number of written research papers itself is small. In Korea, we can observe a rapid rise in the words in 2013. This is because of a sharp increase in research papers from Korean institutes in 2013. Considering this fact, Korea seems to be constantly conducting research that catches up to international research trends. In Japan, it was shown that the words 'social' and 'motion' showed a double increase in 2015 compared to 2011. This opposes the general flow of the research, and it is against the trend in North American and Europe, the continents that lead robotics research. Considering this, we could say that Japan is greatly interested in researching the social function of robots. In North America, we can see that the words 'sensor' and 'interact' constantly increase from 2011 to 2015. It seems that North America focuses on the research that facilitates the sensor-based HRI.

Afterwards, in order to observe the robotics research field in general, we analyzed category frequency. 'Inference' has shown the highest frequency except for 'Others.' Considering that although the words in the top ranks are related to 'Expression,' more words are in the middle and low rank, and we could say that the research on robots inferring human intention or specific situations was the main trend.

Table 1 Documents term matrix

No	autonom	children	gestur	Imag	interact	learn	motion
1	0.071834	0	0.426668	0.071276	0.02811	0	0
2	0	0	0	0	0.08184	0	0.037605
3	0.143668	0.162562	0	0	0	0	0
4	0	0	0	0	0	0.029357	0
5	0	0	0.109037	0.091075	0	0	0.495138
2203	0	0	0	0	0.061575	0	0

Table 2 Top100 keywords on robot research by TF-IDF

No	keyword	TF-IDF	Category	No	keyword	TF-IDF	Category	No	keyword	TF-IDF	Category
1	interact	64.3679	E	36	framework	20.2531	A	71	individu	17.0462	I
2	learn	60.9256	I	37	factor	20.2138	O	72	influenc	16.8701	C
3	social	51.1942	A	38	situat	20.0896	I	73	joint	16.8662	O
4	children	41.6892	A	39	navig	20.0871	I	74	convers	16.8660	E
5	motion	41.3770	E	40	represent	19.9945	E	75	futur	16.5910	O
6	gestur	38.0805	E	41	network	19.6270	A	76	initi	16.5756	O
7	imag	37.6586	E	42	strategi	19.5101	O	77	help	16.5675	I
8	visual	31.9317	E	43	consid	19.3913	I	78	extract	16.4472	I
9	featur	31.8236	I	44	skill	19.3831	A	79	cue	16.3201	I
10	interfac	30.5040	E	45	feedback	19.2188	I	80	platform	16.2375	O
11	sensor	27.4963	C	46	motiv	18.8617	I	81	touch	15.8746	C
12	video	27.0425	I	47	infant	18.6955	A	82	speed	15.8389	O
13	posit	26.4112	I	48	concept	18.6166	O	83	limit	15.5990	O
14	predict	25.9663	I	49	behaviour	18.6103	I	84	guid	15.5802	O
15	gaze	25.5409	E	50	speech	18.5129	E	85	recogn	15.5140	I
16	humanoid	25.2181	O	51	respons	18.4815	I	86	student	15.4642	A
17	movement	24.8777	O	52	teleoper	18.2779	O	87	relationship	15.3584	A
18	autonom	24.8063	O	53	pose	18.1644	E	88	signal	15.2173	I
19	percept	24.2693	C	54	embodi	18.1577	O	89	analyz	15.1621	I
20	camera	23.9150	C	55	virtual	18.0999	I	90	vision	15.1371	E
21	game	23.6885	A	56	role	17.9318	I	91	build	15.0258	O
22	mobil	22.9338	A	57	appear	17.9289	E	92	relat	14.9903	O
23	cognit	22.7553	C	58	walk	17.8941	O	93	teach	14.9635	A
24	space	22.4682	I	59	display	17.6364	E	94	term	14.9509	O
25	pattern	22.0866	I	60	complex	17.5718	O	95	address	14.9507	I
26	collabor	21.8059	A	61	spatial	17.5424	I	96	anim	14.8924	O
27	goal	21.8006	O	62	remot	17.4763	O	97	determin	14.8903	I
28	context	21.1095	C	63	forc	17.4458	O	98	eye	14.8845	I
29	target	21.0889	I	64	knowledg	17.3545	I	99	tool	14.7771	O
30	intent	21.0567	I	65	motor	17.2666	O	100	distanc	14.7769	I
31	bodi	20.9753	E	66	report	17.2589	O				
32	languag	20.9256	E	67	architectur	17.2532	O				
33	mechan	20.8476	O	68	identifi	17.1852	I				
34	facial	20.5720	E	69	project	17.0901	O				
35	devic	20.3127	O	70	program	17.0605	O				

C=Cognition, **E**=Expression, **I**=Inference, **A**=Application, **O**=Others

Figure 3 Top 10 emerging keywords in 2015 compared to 2011

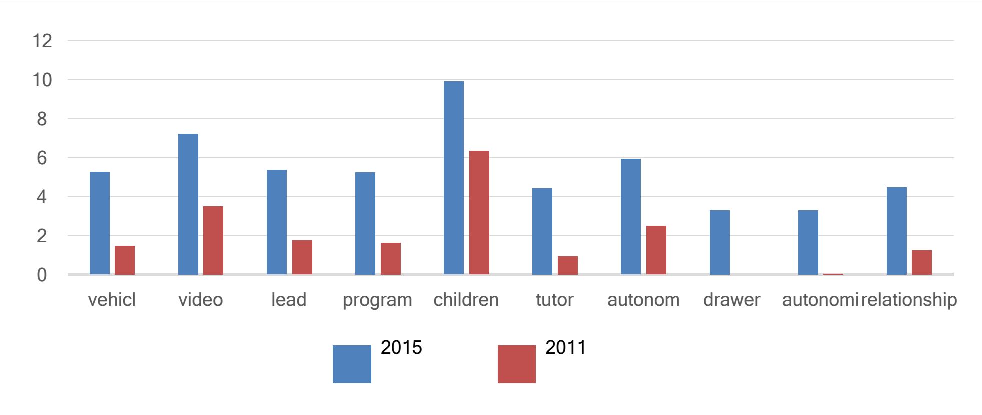


Figure 4 Top 10 disappearing keywords in 2015 compared to 2011

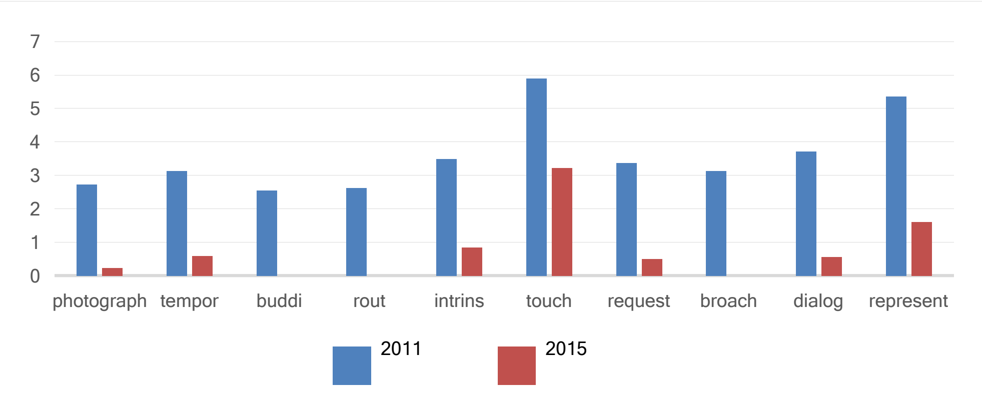
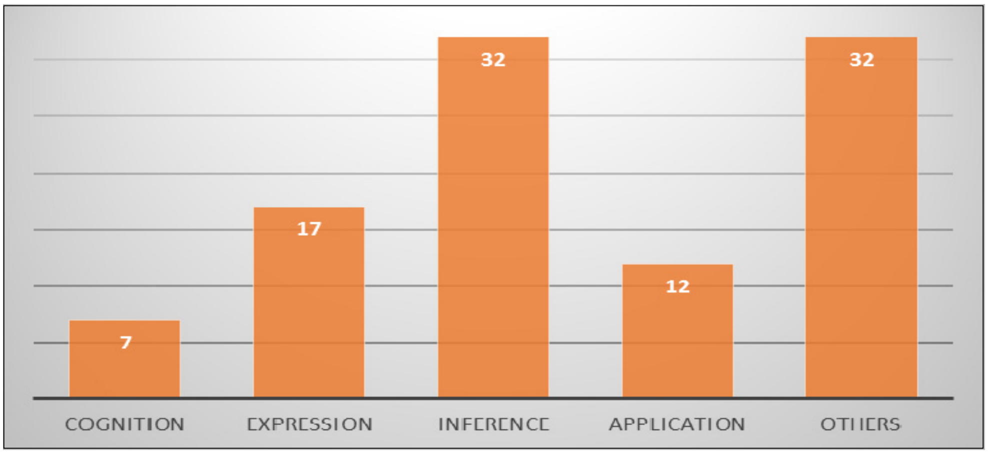


Figure 5 Frequency by robotic category

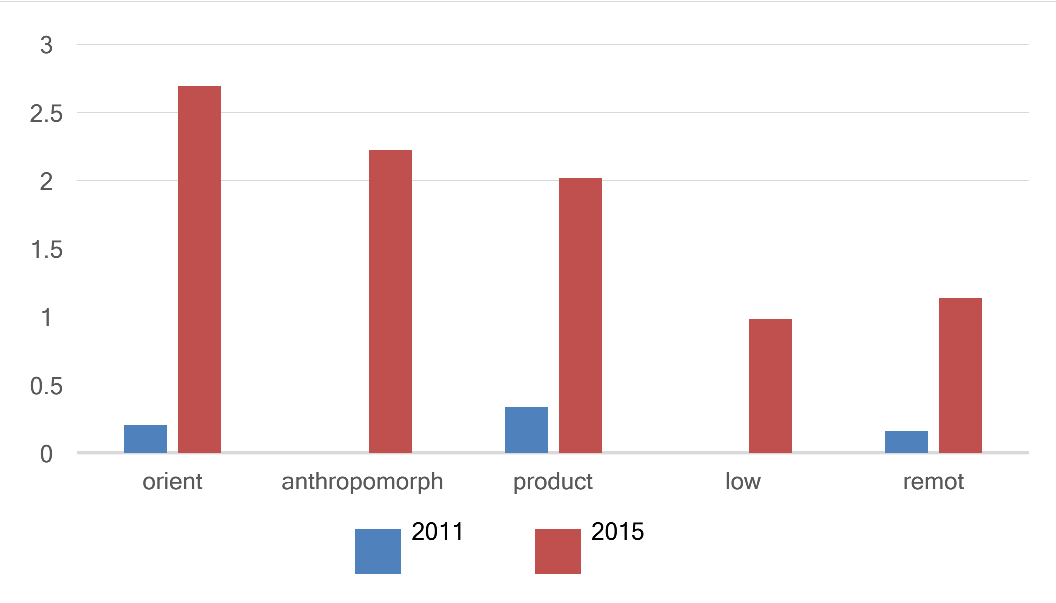


Analysis on research fields that showed increase in 2015 compared to 2011 by continent

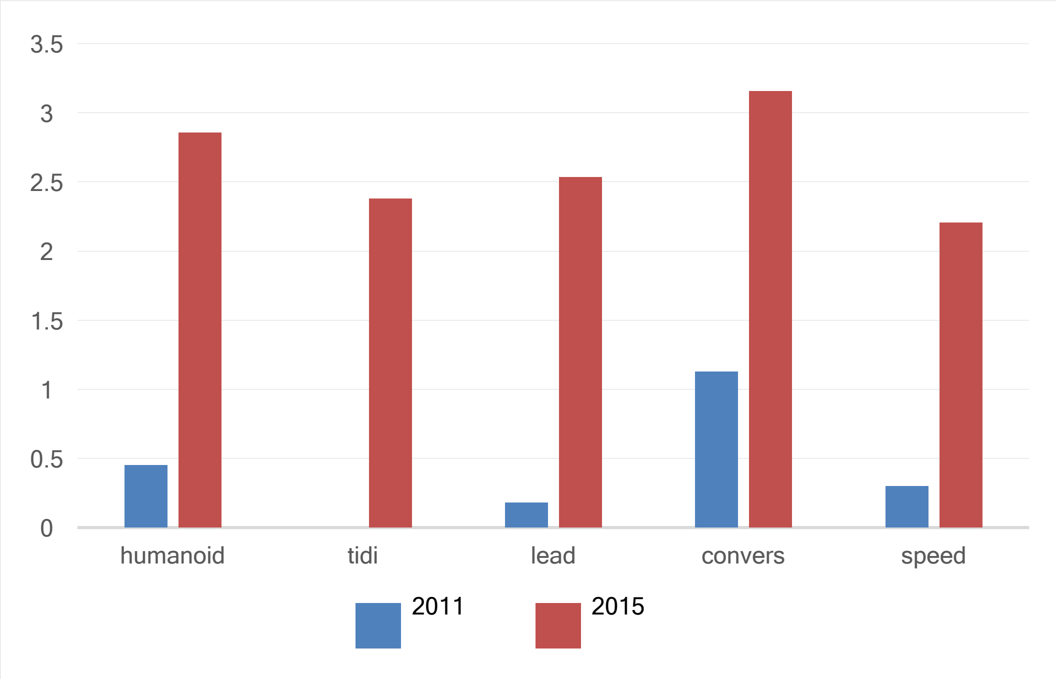
In order to find new research frontiers by continent (Korea, Japan, North America, EU), we selected 5 top keywords that showed the largest TF-IDF value difference in 2015 to 2011 as hot keywords and compared them. But since Asian countries other than Korea, China, and Japan and South America and Africa have very few research papers, and with China, there was the problem of selecting hot keywords because there were few constant keywords from 2011 to 2015, they were excluded from this analysis. In this way, we selected hot keywords that showed the largest increases in TF-IDF value in 2015 compared to 2011 to explore the new frontier research fields by continent (EU, North America, Korea, Japan).

In the following graphs, Korea showed 'anthropomorph' and Japan showed 'humanoid.' Although different words appeared, they commonly describe humans. Considering this fact, we can see that Korea and Japan are conducting research on the external appearance of robots. Especially for Japan, considering that words 'convers' and 'lead' appeared, it seems that the research on the social robot emerged as a new issue. In Europe, it seems that the research on robots with social function is conducted actively. In North America, unlike Korea, Japan, or Europe, we could see a focus on technological aspects.

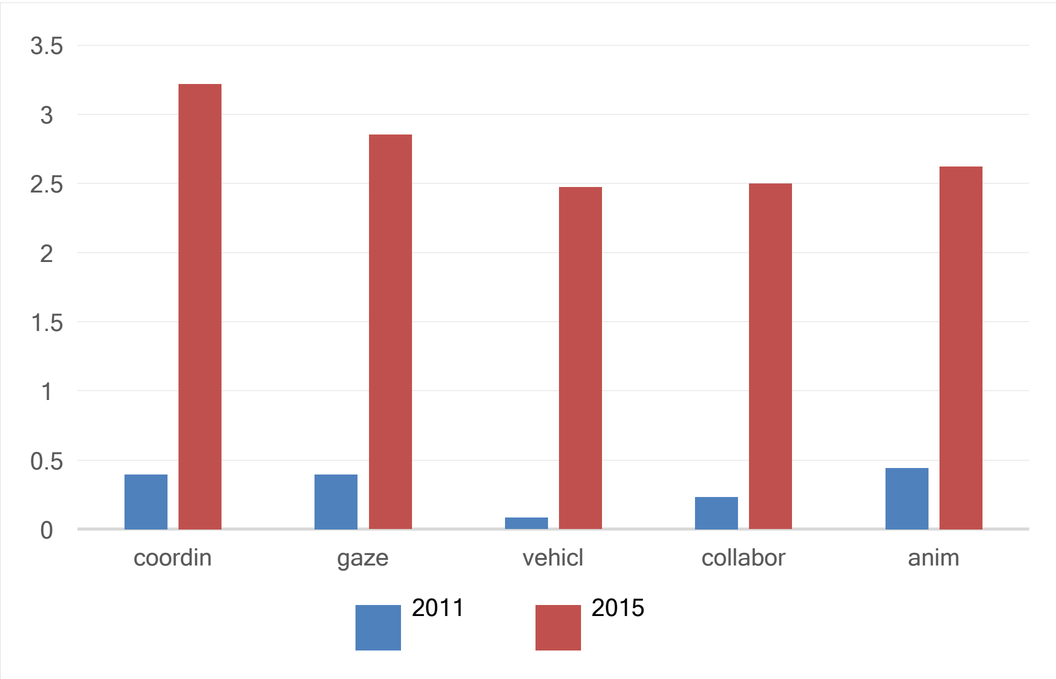
Figure 6 Top 5 keywords in increase in 2015 compared to 2011



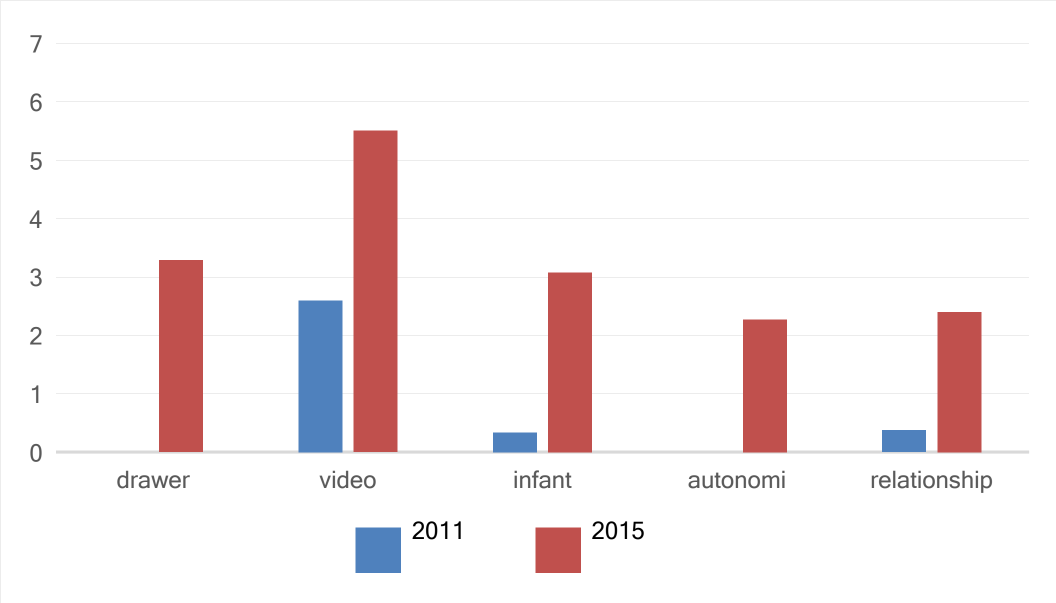
(a) Korea



(b) Japan



(c) Europe



(d) North America

3.3 Keyword Network Analysis

3.3.1 Keyword network by year

We extracted keywords by year to observe the changes in the robotics research trend. We presented the top 30 keywords with the highest TF-IDF values from 2011 to 2015 in network form.

The network appeared to be as in Figure 8(a), and the centrality appeared to be as in Table 5(a). In 2011, we were able to find a cluster around 'extract' (closeness=1.8465115), 'cognit' (closeness=1.845833), 'visual' (closeness=1.8255119), 'touch' (closeness=1.7689986), 'interfac' (closeness=1.7643592), and 'skill' (closeness=1.7164686), with keyword 'percept'(closeness= 1.8951205) at the center. We could see that research with a pivot in 'percept,' which has the highest closeness score, is the mainstream

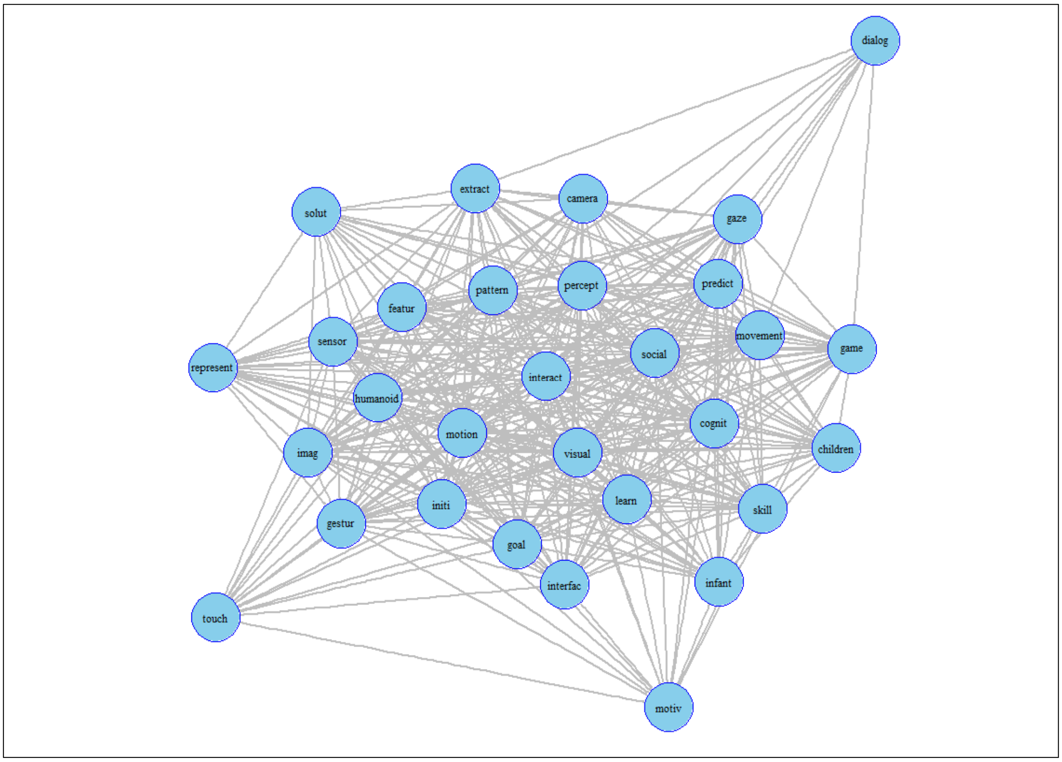
The centrality result in 2012 was as Table 5(b), and many words were connected with 'anim' (closeness=2.38461471) at the center. Terms close to 'anim' were 'space' (closeness=2.37346922), 'facial' (closeness=2.28141316), 'virtual' (closeness=2.16483994), 'pose' (closeness=2.15844241), 'motor' (closeness=2.13865509), 'imag' (closeness=2.02874524), and 'camera' (closeness= 2.00549876). While research regarding perception and cognition was mainstream in 2011, research relatively more focused on movements was mainstream in 2012. Network was as Figure 8(b).

Considering Figure 8(c) and Table 5(c), it seems integrated research of research in 2011 and 2012 were conducted in 2013. With 'color' (closeness=2.006179) at the center, 'mobil'(closeness=1.9011), 'remot'(closeness=1.878968), 'pattern'(closeness=1.756389), 'gestur'(closeness=1.745825), 'virtual' (closeness=1.707528), 'sensor'(closeness=1.696708), 'percept'(closeness=1.693534), 'factor'(closeness =1.683232), 'predict'(closeness=1.630893), and 'video'(closeness=1.616541) are located. Combining the words mentioned previously, we can find that while research on simple pattern recognition or prediction by data analysis through the sensors was conducted in the past, research on learning data collected through the sensor and a patternization of the collected data for prediction were conducted in 2013.

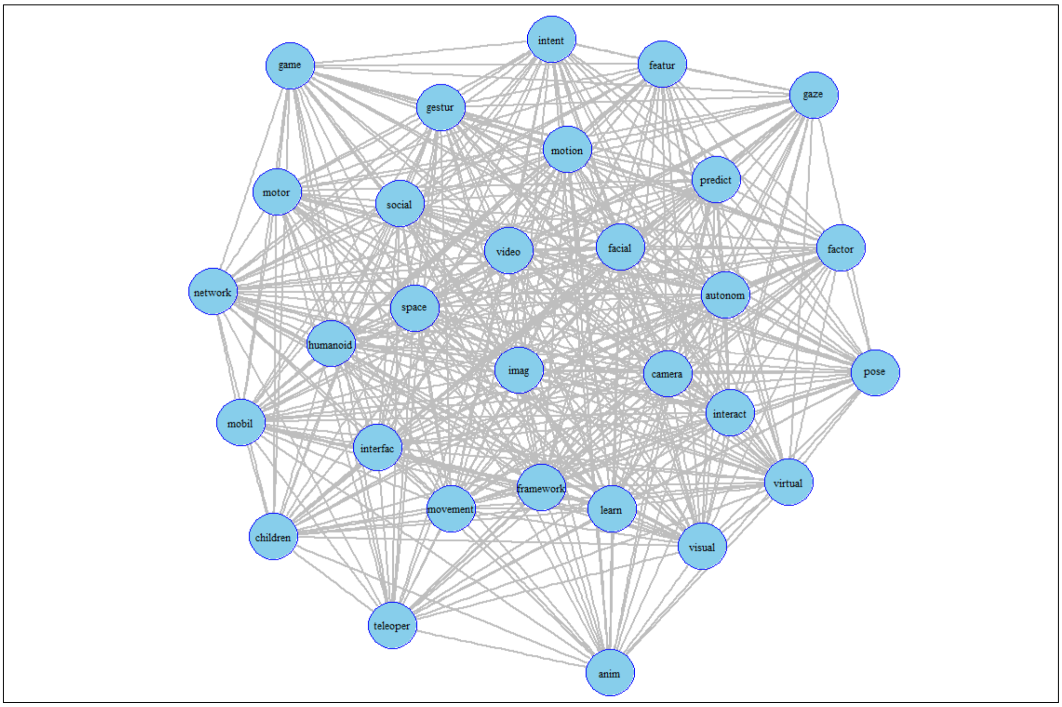
We could find that research that breaks out of simply precisely performing tasks through robotics technology development and studies about robots that could provide services from the users' POV were conducted in 2014, as shown in Figure 8(d).

Lastly, based on Figure 8(e), it seems that the research that improves the previously researched aspects was conducted in 2015, the most recent period.

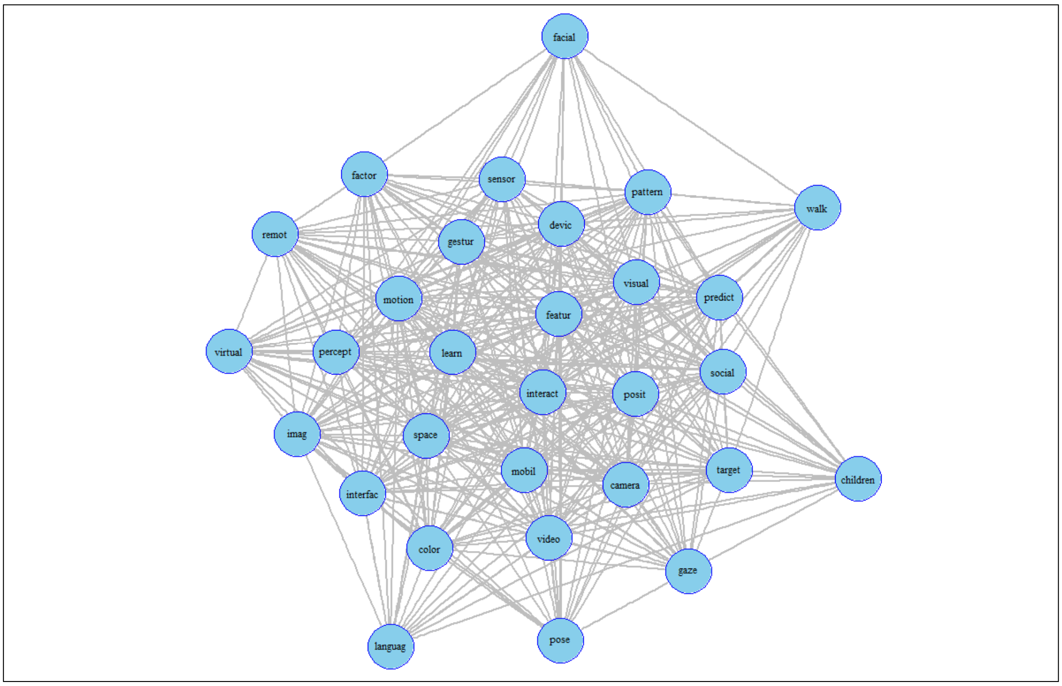
Figure 6 Network analysis by year



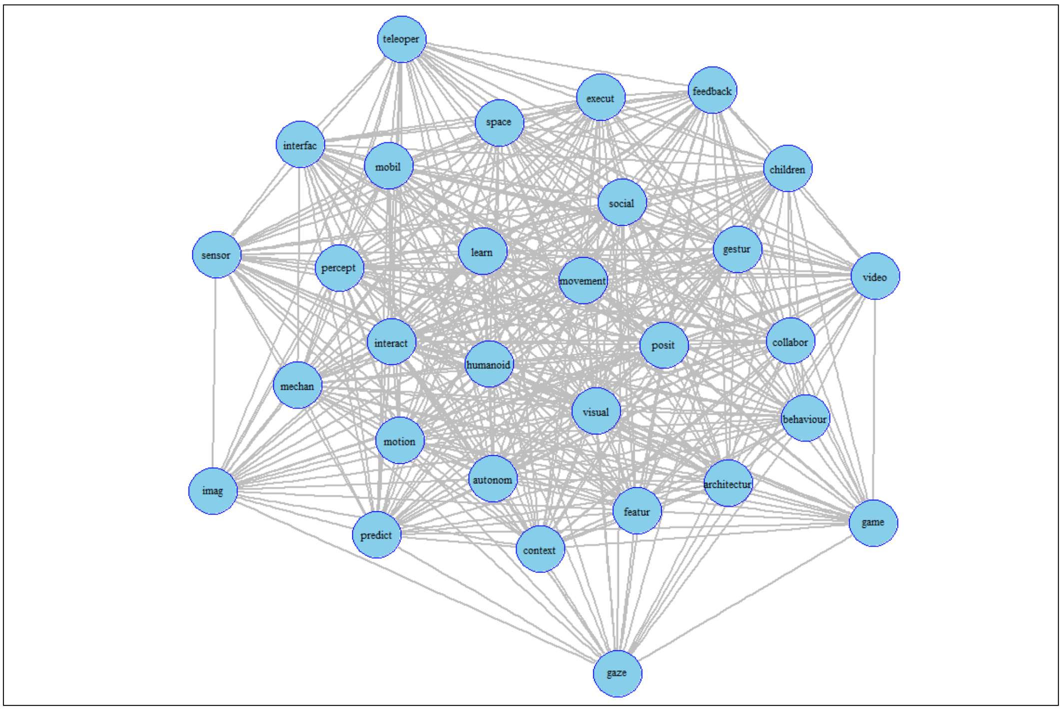
(a) 2011


intent
game
featur
gaze
gestur
motion
predict
motor
social
video
facial
factor
network
space
autonom
humanoid
imag
camera
pose
interact
mobil
interfac
virtual
framework
learn
movement
children
visual
teleoper
anim

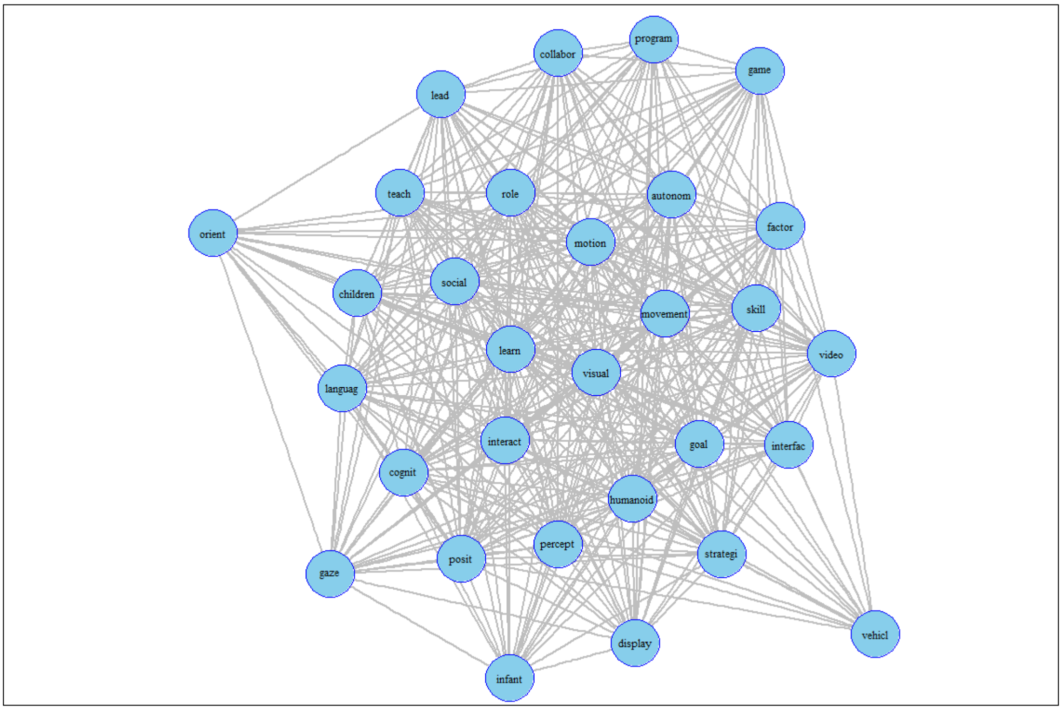
(b) 2012


facial
factor
sensor
pattern
walk
remot
devic
gestur
visual
predict
motion
featur
virtual
percept
learn
social
interact
posit
imag
space
target
mobil
camera
children
interfac
video
color
gaze
languag
pose

(c) 2013


teleoper
execut
feedback
space
interfac
mobil
children
social
sensor
percept
learn
gestur
movement
video
interact
posit
collabor
humanoid
mechan
visual
behaviour
motion
imag
autonom
architectur
featur
game
predict
context
gaze

(d)2014


program
collabor
game
lead
teach
role
autonom
orient
factor
motion
social
children
movement
skill
learn
video
visual
languag
interact
goal
interfac
cognit
humanoid
percept
strategi
gaze
posit
display
vehicl
infant

(e) 2015

Table 3 Network analysis centrality result by year (a) 2011~2012 network analysis centrality result

2011				2012			
term	**Betweenness**	**Closeness**	**Degree**	**term**	**Betweenness**	**Closeness**	**Degree**
percept	69	1.8951205	28	anim	91	2.38461471	22
extract	54	1.8465115	24	space	76	2.37346922	29
cognit	64	1.845833	27	facial	34	2.28141316	28
visual	34	1.8255119	28	virtual	9	2.16483994	28
touch	52	1.7689986	15	pose	54	2.15844241	26
interfac	40	1.7643592	25	motor	63	2.13865509	26
skill	43	1.7164686	25	imag	29	2.02874524	29
solut	55	1.6964985	19	camera	21	2.00549876	26
motion	15	1.6784436	27	intent	33	1.99219438	26
goal	21	1.6596656	25	game	47	1.96390656	23
pattern	32	1.6583858	26	factor	73	1.95655129	25
sensor	23	1.6555287	27	network	20	1.94637943	26
represent	19	1.6362334	21	children	7	1.94453393	24
gestur	22	1.5314873	24	video	50	1.93918253	29
camera	28	1.5274583	22	movement	41	1.88590786	29
imag	5	1.524971	26	visual	28	1.87685673	28
featur	17	1.518547	27	autonom	5	1.8394565	29
initi	0	1.3136824	28	framework	33	1.79832607	27
game	6	1.3075227	25	featur	0	1.64462531	26
gaze	3	1.2891956	23	interfac	13	1.58858454	28
movement	0	1.2836734	26	gaze	0	1.581909	23
infant	2	1.2828067	21	social	3	1.49272899	29
humanoid	0	1.2767471	27	predict	0	1.4725183	28
predict	0	1.2731087	27	mobil	0	1.44944828	27
social	0	1.1952782	28	teleoper	5	1.43282873	24
children	0	1.087898	22	learn	0	1.3812839	29
motiv	0	1.079346	15	humanoid	0	1.29028805	29
interact	0	0.9973744	29	gestur	0	1.15963331	28
dialog	0	0.9543796	9	motion	0	0.67897035	28
learn	0	0.9453163	28	interact	0	0.54842716	29

Table 3 Network analysis centrality result by year _ (b) 2013~2014 network analysis centrality result

2013				2014			
term	**Betweenness**	**Closeness**	**Degree**	**term**	**Betweenness**	**Closeness**	**Degree**
color	141	2.006179	24	interfac	121	1.984668	25
mobil	51	1.9011	28	social	34	1.926595	29
remot	72	1.878968	21	context	105	1.851813	27
pattern	38	1.756389	26	mechan	89	1.831469	28
gestur	34	1.745825	26	mobil	0	1.786482	26
virtual	21	1.707528	20	architectur	56	1.731435	27
sensor	46	1.696708	26	gestur	16	1.612423	28
percept	0	1.693534	26	humanoid	15	1.59522	28
factor	61	1.683232	22	game	63	1.581722	22
predict	70	1.630893	27	space	13	1.573374	27
video	42	1.616541	27	movement	38	1.565042	29
learn	1	1.593923	27	imag	84	1.546986	21
camera	29	1.548858	27	feedback	19	1.521907	26
languag	35	1.546081	17	posit	6	1.513819	29
social	27	1.525999	26	teleoper	35	1.503349	23
posit	18	1.515075	27	visual	2	1.498371	29
devic	7	1.504468	26	collabor	14	1.450152	27
interfac	4	1.49825	26	sensor	12	1.344039	25
interact	0	1.376115	29	children	0	1.313311	26
space	0	1.288661	27	behaviour	0	1.301287	27
motion	9	1.237948	28	featur	0	1.250942	28
target	2	1.213076	24	percept	3	1.242008	28
gaze	4	1.207485	22	motion	0	1.178129	29
children	0	1.15856	17	video	0	1.173957	24
facial	0	1.060708	16	predict	0	1.107426	27
walk	0	1.056349	16	execut	0	1.030847	26
imag	0	1.046584	25	autonom	0	0.894696	29
pose	0	1.00582	18	gaze	0	0.824983	20
visual	0	0.936905	27	learn	0	0.765534	29
featur	0	0.813801	28	interact	0	0.451434	29

Table 3 Network analysis centrality result by year _ (c) 2015 network analysis centrality result

2015				2015			
term	**Betweenness**	**Closeness**	**Degree**	**term**	**Betweenness**	**Closeness**	**Degree**
strategi	153	2.746371	26	role	9	1.870889	28
display	147	2.696866	23	humanoid	9	1.864156	27
motion	31	2.67529	28	video	5	1.85042	26
teach	25	2.359967	24	gaze	11	1.801616	22
lead	43	2.298251	25	languag	0	1.735179	25
percept	35	2.283634	26	posit	11	1.726165	26
collabor	16	2.106109	23	social	0	1.716121	28
learn	44	2.039002	29	autonom	4	1.680699	28
factor	37	1.996634	26	vehicl	33	1.644553	15
movement	27	1.982297	28	game	0	1.505231	21
program	27	1.973163	22	cognit	8	1.480466	26
goal	15	1.960564	27	skill	0	1.450198	26
infant	46	1.960508	21	visual	1	1.405171	29
orient	13	1.945449	16	children	0	1.060473	26
interfac	19	1.931712	26	interact	0	0.690714	29

3.3.1 Keyword network by continent

Table 6(a) is the centrality result in network analysis on Korea and Japan. In Korea, considering that 'target' marked the highest closeness by 1.017801, the word is considered to be at the center of the words, and the words 'intellig'(closeness=0.947407), 'motion'(closeness=0.912582), and 'sensor'(closeness=0.878567) are close to the center. Considering the words gathering around 'target,' Korea seems mainly to conduct research of recognizing the target human or object movement through sensors. Also, in Japan, the word 'game' (closeness=1.97052) is the center-most word, but 'game' and 'target' (closeness=1.537418) are in very close together. Japan has also shown similar words in the center of the network. Considering this, we can find that Japan and Korea conduct similar research. Next, Table 6(b) shows the centrality in network analysis in Asia countries other than Korea, China, and Japan, and China. Considering the closeness of the two results, it was found that there are less differences between the words. However, in China, we can see the rapid fall in closeness in words like 'gestur,' 'wave,' 'forearm,' 'handpuppet,' and 'calibr.'

If we take a look at the European word network structure through Table 6(c), 'gestur' (closeness=0.569418) is the center and the words 'tutor' (closeness=0.543766), 'skill' (closeness=0.539571), 'motor' (closeness=0.538629), and 'posit' (closeness=0.506314) gather around the center. Next, if we take a look at the closeness in North America shown in Table 6(c), we can say 'predict,' with the highest closeness by 1.20077, is in the center of the network.

Table 4 Network analysis centrality result by continent _ (a) Korea, Japan network analysis centrality result

Korea				Japan			
term	Betweenness	Closeness	Degree	term	Betweenness	Closeness	Degree
target	179	1.017801	14	game	109	1.597052	24
intellig	111	0.947407	22	target	111	1.537418	25
orient	67	0.915877	18	devic	59	1.439421	27
motion	77	0.912582	15	learn	43	1.388807	26
sensor	81	0.878567	20	pattern	76	1.385871	24
social	61	0.877264	20	embodi	34	1.338059	21
quadrotor	0	0.828512	6	eye	33	1.330706	17
devic	32	0.815715	12	convers	9	1.323293	24
match	48	0.792926	8	posit	25	1.305189	28
interact	101	0.768053	28	elder	44	1.301027	24
posit	18	0.758572	25	virtual	30	1.2788	24
children	21	0.753209	12	walk	0	1.241515	21
intent	32	0.739916	17	speech	32	1.186289	23
percept	1	0.732337	15	shop	12	1.18429	20
teleoper	0	0.717206	10	util	38	1.175449	28
pattern	18	0.716476	14	touch	13	1.158479	23
smart	9	0.692241	9	gaze	21	1.1149	23
anthropomorph	0	0.678809	9	motion	3	1.10549	28
interview	2	0.662093	6	consid	1	1.066786	28
product	0	0.587132	12	factor	4	1.052	25
productori	0	0.585008	11	bodi	2	1.031052	28
camera	0	0.510954	14	result	0	0.991959	27
humanori	0	0.503696	11	social	0	0.983835	28
appear	0	0.47629	14	interfac	0	0.932235	29
a	0	0.471516	16	children	3	0.925462	19
pose	0	0.466286	7	situat	3	0.923343	26
ride	0	0.41952	6	confirm	0	0.827276	29
implic	0	0.419195	17	humanoid	0	0.769738	25
gestur	0	0.354632	10	wheelchair	0	0.493956	10
hors	0	0.326119	4	interact	0	0.472123	28

Table 4 Network analysis centrality result by continent _ (b) Asia other, China network analysis centrality result

Asia other				China			
term	Betweenness	Closeness	Degree	term	Betweenness	Closeness	Degree
spatial	186	0.031607	11	learn	183	0.006506	11
consid	192	0.031599	17	visual	146	0.006506	10
learn	78	0.031575	17	network	80	0.006503	6
interact	116	0.031557	24	sensori	0	0.006496	10
pattern	0	0.031547	9	vergenc	39	0.006492	8
autist	48	0.031483	7	symbollik	0	0.006492	4
product	0	0.031449	4	optim	135	0.006488	9
respons	17	0.031449	9	reward	30	0.006485	8
kinemat	0	0.031437	5	represent	0	0.006478	10
children	2	0.031424	10	polici	0	0.006473	8
camera	51	0.031405	8	eigenspac	0	0.006466	3
social	14	0.031402	18	reidentif	0	0.006466	3
autism	0	0.031352	8	dispar	0	0.006458	7
motion	33	0.03128	11	climb	0	0.006445	4
percept	54	0.03101	16	wall	0	0.006443	4
screen	27	0.030767	6	binocular	0	0.006436	8
music	0	0.030695	4	electrostat	0	0.006433	4
exercis	0	0.030614	5	rfid	0	0.006408	1
polit	0	0.0304	4	adhes	0	0.006407	4
grasp	27	0.030349	3	rout	45	0.006274	4
anim	0	0.029973	6	door	0	0.006189	2
workshop	0	0.029898	7	corner	0	0.006183	2
voltag	0	0.02955	1	workspac	0	0.006171	1
gender	0	0.029056	5	dtn	0	0.006009	2
maxim	0	0.027971	3	predictor	0	0.005966	1
aoa	0	0.027528	1	gestur	0	0.00128	3
robowait	0	0.027418	3	wave	0	0.001279	3
student	0	0.027287	9	forearm	0	0.001279	3
media	0	0.027179	3	handpuppet	0	0.001279	3
moral	0	0.001149	0	calibr	0	0.001149	0

Table 4 Network analysis centrality result by continent _ (c) Europe, North America network analysis centrality result

Europe				North America			
term	Betweenness	Closeness	Degree	term	Betweenness	Closeness	Degree
gestur	117	0.569418	28	predict	80	1.200777	29
tutor	48	0.543766	20	command	79	1.176266	25
skill	13	0.539571	29	gaze	68	1.146821	26
motor	60	0.538629	28	path	77	1.146148	22
posit	51	0.506314	28	remot	63	1.145637	27
strategi	30	0.498779	29	video	51	1.143382	29
predict	9	0.487006	28	sensor	42	1.067351	28
motiv	32	0.484454	29	trust	12	1.016444	13
movement	22	0.472706	28	pose	21	0.976692	29
architectur	17	0.461679	28	collabor	35	0.948066	26
context	2	0.445764	28	spatial	49	0.891574	27
children	46	0.440964	29	represent	25	0.878612	27
project	8	0.4377	28	imag	29	0.872901	25
motion	11	0.429457	27	children	5	0.837389	25
game	15	0.428643	28	motion	2	0.836408	28
feedback	17	0.426974	29	learn	0	0.807725	29
space	23	0.423718	29	infant	16	0.791957	20
sensorimotor	7	0.417934	24	interfac	22	0.767436	26
autonom	0	0.417381	29	camera	5	0.765743	26
complex	7	0.414873	29	gestur	20	0.746536	26
cognit	0	0.401909	28	vision	3	0.745019	26
visual	13	0.401697	29	featur	3	0.708362	27
percept	2	0.393538	29	interact	0	0.692569	29
mechan	0	0.391651	29	languag	0	0.692157	29
goal	5	0.378308	29	social	0	0.688261	25
behaviour	0	0.361695	29	autonom	3	0.688054	28
social	0	0.295294	29	search	6	0.684365	28
humanoid	0	0.287364	29	framework	2	0.665914	28
learn	0	0.190276	29	visual	0	0.654477	28
interact	0	0.15593	29	mobil	5	0.624559	27

4. Discussion

This study compared the topics of robotics research through the abstracts of conference papers from regions around the world from 2011 to 2015. First, we were able to find that current research focuses on interaction, that research on motions or images is also actively conducted, and that recognition technology through sensors is frequently studied.

Second, although there was no significant difference in the utilization fields of the robotics research in each region, the core research subjects by country or region are diverse. Hence, in order to establish an exemplary success in the common utilization field as interest in robotics development around the world rises, establishing a converged network or consortium in which entities from various countries participate to facilitate exchanges among research units would be useful. Especially in the case of China, supply is obviously insufficient in contrast to the enormous robot market demand, and although Chinese industrial robotics research significantly improved because the demand in industrial robots is especially great, China is only in the early stages of research results (Gao et.al, 2015). Hence, development in the HRI field is delayed, and balanced research is difficult. Therefore, China should make internal efforts to encourage market competition in a beneficial direction through public research, following the suggestion of Cohen et al. Externally, if China encourages forming a network with neighboring Korea and Japan to converge various industries and robotics technology, patent accumulation from enhanced R&D would naturally lead to improved industrial competitiveness.

Third, we found that academia conducts a great deal of research related to interaction between humans and robots, i.e., robot sociality. In light of this fact, humanoid may sometime in the near future. The possibility of invigoration of the robot industry to not only simply enhance industrial competitiveness but also to solve social issues is increasing. Recently, the supply of caregivers who could help elderly people with diseases is not meeting the increase of the elderly population (Tang et al, 2015). However, if the sociality of robots continues to develop, we may expect to see some help in addressing the problems of an aging society. Also, considering the previously deducted results, it was found that the research on rehabilitation robot is insufficient. Hence, if we invigorate the research on rehabilitation robots, we may become able to support not only the elderly but also a wider range of socially disadvantaged classes.

Next, the research on HRI is actively conducted. However, as most of the robotics researchers have been proceeding their studies in accordance with the engineering curricula, they lack the liberal arts aspects in engineering (Hynes and Swenson, 2013). Although they aim for a human-robot user-friendly interface, the actual application might be difficult since the researches have usually lacked a background in the liberal arts. Therefore, it is expected that research results with better quality would be deducted if the curriculum that fosters the liberal arts refinement is more established or actively converged research between robotics researchers and liberal arts researchers is conducted.

Also, exploring the Appendix, we found that the papers differ by the venues of the conference. It is concluded that this is because of the physical distance between the institute and the conference venue and the high cost of transport. Such aspects may hinder the development of robotics research. It is because researchers lose the opportunity to present their discoveries due to financial and time costs although they have important research performances. Therefore, active efforts such as remitting the conference participation cost of the researcher from a distant region or providing the opportunity to present on-line without actually attending the conference venue through actively utilizing IT technologies would be needed.

The result of this study is based on the contents of papers presented in major robotics conferences from 2011 to 2015. The reason for not choosing an academic journal is that the presenting time of journals is later than that of conferences so that the journals may not reflect the reality as sensitively as the conferences, but there is need for future analyses with more diverse data.

The text mining method for the analysis of unstructured big data is gathering attention for its objectivity, shorter analysis time, and cost-efficient empirical testing. We can find trends in robotics research through the text mining method and can secure several meaningful points. Since on-line analysis is possible whenever robotics research materials are updated, we would be able to constantly monitor the changes in trends and issues.

Received 09. Jan. 2017, Accepted 27. Feb. 2017

References

Cohen, W. M., R. R. Nelson, and J. P. Walsh., 2002, Links and impacts: the influence of public research on industrial R&D, *Management science* 48, 1-23.

Gao, H., R. Guo, and P. Li, 2015, Development Situation and Prospect of Chinese Industrial Robots, *2nd International Conference on Civil,* Materials and Environmental Sciences, Atlantis Press.

Griffiths, T. L. and M. Steyvers, 2004, Finding scientific topics, *Proceedings of the National Academy of Sciences* 101, 5228-5235.

Hynes, M. and J. Swenson, 2013, The humanistic side of engineering: considering social science and humanities dimensions of engineering in education and research, *Journal of Pre-College Engineering Education Research (J-PEER)* 3, 31-42.

Kang, M. K., H. J Lee, and Y. T. Kim, 2013, The current status of the use of robots for children with disabilities, *Special Education Research* 12, 127-152.

Kim, S. B. and Y. H Sin, 2014, A review of studies on the application of robots for improvement of social skills for children with autistic spectrum disorder and emotional and behavioral disorder, *Journal of Special Education & Rehabilitation Science* 53, 21-52.

Kim, C., 2012, An analysis of domestic research trend and educational effects in relation to robot education, *Korean Association of Information Education* 16, 233-243.

Lee, S. B. and K. Jeong, 2015, Research trends in robotics, control, and automation based on 2011 ~ 2014 IEEE ICRA Proceedings (Part 2). *Journal of the Korean Society for Precision Engineering* 32, 317-325.

Lee, H. M, J. B. Kim and S. H Kim, 2013, Research trends and issue of robot teaching technology, *Institute of Control, Robotics and Systems,* 19, 49-59.

Lechevalier, S., Y. Ikeda, and J. Nishimura, 2011, Investigating collaborative R&D using patent data: the case study of robot technology in Japan, *Managerial and Decision Economics* 32, 305-323.

Lechevalier, S., Y. Ikeda, and J. Nishimura, 2010, The effect of participation in government consortia on the R&D productivity of firms: a case study of robot technology in Japan, *Economics of Innovation and New Technology* 19, 669-692.

Park, Y. J., M. G, Sin, J. W, Kim, and K. J. Cho, 2013, Research trends of biomedical & rehabilitation Robots, *Korea Robotics Society* 10, 27-32.

Song, M. and S. Y. Kim, 2013, Detecting the knowledge structure of bioinformatics by mining full-text collections, *Scientometrics* 96, 183-201.

Tang, D., B. Yusuf, J. Botzheim, and N. Kubota, 2015, A novel multimodal communication framework using robot partner for aging population, *Expert Systems with Applications* 42, 4540-4555.

Worldrobotics.org [homepage on the Internet]. International Federation of Robotics, Frankfurt [updated 12 October 2016; cited 2015]. Available from: http://www.worldrobotics.org/

Appendix A. Conference materials used in survey and analysis

Table 1 ARSO

	2011	2012	2013	2014	2015
Venue	California, USA	Munich, Germany	Tokyo, Japan	Illinois, USA	Lyon, France
South Korea	2	1	0	0	0
China	1	0	1	2	1
Japan	4	4	30	5	5
Asia(Others)	0	0	3	1	1
Europe	8	15	7	2	21
North America	5	1	0	15	0
Africa	0	0	1	0	0
Oceania	1	0	0	0	0
South America	0	0	0	0	1

Table 2 CRV

	2011	2012	2013	2014	2015
Venue	Newfoundland, Canada	Ontario, Canada	Saskatchewan, Canada	Quebec, Canada	Nova Scotia, Canada
South Korea	0	1	0	0	0
China	1	1	0	0	0
Japan	0	1	0	0	0
Asia(Others)	4	4	1	0	0
Europe	4	9	4	9	7
Notrh America	41	52	44	41	36
Africa	1	0	0	2	1
Oceania	2	0	1	0	0
South America	0	0	2	0	0

Table 3 HRI

	2011	2012	2013	2014	2015
Venue	Lausanne, Switzerland	Massachusetts, USA	Tokyo, Japan	Bielefeld, Germany	Portland, USA
South Korea	13	9	10	9	14
China	0	0	0	0	0
Japan	21	21	34	17	17
Asia(Others)	3	7	4	8	0
Europe	29	34	36	77	41
Notrh America	35	66	37	38	75
Africa	1	0	2	0	0
Oceania	2	5	5	4	0
South America	0	3	2	2	2

Table 4 ICDL

	2011	2012	2013	2014	2015
Venue	Frankfurt, Germany	California, USA	Osaka, Japan	Genoa, Italy	Rhode Island, USA
South Korea	0	0	1	1	1
China	1	3	0	0	2
Japan	5	8	11	12	6
Asia(Others)	1	3	0	1	1
Europe	45	45	31	64	31
Notrh America	15	39	10	5	19
Africa	0	1	0	1	1
Oceania	0	2	0	1	0
South America	0	0	0	0	0

Table 5 RO – MAN

	2011	2012	2013	2014	2015
Venue	Atlanta, USA	Paris, France	Gyeongju, Korea	Edinburgh, UK	Kobe, JAPAN
South Korea	4	6	54	11	3
China	0	0	1	2	1
Japan	23	69	43	35	59
Asia(Others)	0	7	6	6	6
Europe	28	69	47	88	40
Notrh America	29	21	14	34	21
Africa	0	0	1	0	0
Oceania	0	0	1	5	6
South America	0	1	1	3	1