Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

A Bibliometric Approach to Multimodal Interaction Research Trends

Jianan Lyu[1,2,*], Abdul Aziz Zalay[1], Nor Syazwani Binti Mat Salleh[1]
[1]Faculty of Art, Computing and Creative Industry, Sultan Idris Education University, 35900, Malaysia
[2]School of Art and Design, Huizhou University, Huizhou 516007, China
*Corresponding Author.

Abstract:

The demand for faster, more efficient, and higher recognition of human-computer interactions, especially for critical task completion rates, highlights the importance of multimodal interactions. This paper aims to identify global trends in multimodal interactions and their most relevant research areas. A total of 401 papers published in the Scopus database from 2012 to 2021 were selected for bibliometric analysis. We refined all Scopus categories related to multimodal interaction to obtain bibliometric information. This paper provides a qualitative analysis of the first 20 cited documents. Bibliometric analysis during the study revealed the number of publications per year, the growth rate of producing and cooperating countries, and sources. The United States ranked first in paper production. Publications after 2016 focus more on "visualization," "augmented reality." Computer Science Lectures (Including Artificial Intelligence Sub-series Lectures and Bioinformatics Lectures) is the most representative journal in the multimodal interaction research area. Keywords such as multimodal interaction and multimodality are increasing their visibility. "Multimodal interactions," "interactive computer systems," and "user interface" are more likely to be used in future research on multimodal interaction.

Keywords: *Academics, Bibliometric analysis, Multimodal interaction, Human-computer interaction, Trends*

I. INTRODUCTION

"Modality" is a biological concept proposed by German philosopher Helmholtz, which refers to the channels through which organisms receive information through sensory organs and experience [1]. For instance, humans have vision, hearing, taste, touch, and smell. Face recognition and behavior detection in vision, speech recognition and speech synthesis speech, and human-machine dialogue in natural language. These fields were previously independently evolving technologies and applications, which can be considered as single-modal technologies. Different modal forms describe the characteristics of the same object from different angles.

Multimodal interaction includes sensory interactions such as vision, hearing, smell, touch, and taste, realized through the eyes, ears, nose, mouth, and skin touch. The application of technology in real life is

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

also designed around these sensors. Combine the interactive technologies of multiple senses to form a multimodal interactive form.

The essence of multi-modality is that the multi-modal signals of different channels complement each other, discovering detailed features or feature combinations, which is helpful for the expansion of artificial intelligence (AI) application scenarios. Artificial intelligence is the imitation and learning of humans, and humans are the intelligent body that sees, listen, and speaks multimodally work together. So, the future development direction of artificial intelligence applications is multimodal interaction technology [2].

To identify gaps in multimodal interaction research in various areas, researchers must explore or review sources and databases of published documents. The Scopus database is the world's largest abstract and citation database, covering 15,000 scientific, technical, and medical journals [3]. Meanwhile, Scopus is the most complete and famous online scientific citation search tool [4,5]. Therefore, it is a credible reference for researchers publishing and seeking the recent technologies, challenges, trends, experiments, enhancements, and research opportunities.

Bibliometrics is an interdisciplinary science, which uses mathematical and statistical methods to analyze all knowledge carriers quantitatively [6]. It is a comprehensive knowledge system, integrating mathematics, statistics, and philology, emphasizing quantification. Bibliometric methods are often used to assess scientific manuscripts to determine research tendencies [5,7,8]. The main assessment objects are the number of literature (different publications, especially journal papers and citations), the number of authors (individual collectives or groups), and vocabulary (various literature identifiers, most of which are thesaurus). Bibliometrics has been applied to assess scientific advances in many scientific and engineering disciplines. It is a common research tool for publication's systematic analysis [9-13]. This paper applied the bibliometric analysis method to the "multimodal interaction" area.

This study used bibliometric tools to identify the main and significant research trends in multimodal interaction studies, as well as the most related research areas in which multimodal interactions have a substantial impact. Bibliometric information about "multi-modal interactions" was extracted from Elsevier's Scopus database.

Scopus delivers an interdisciplinary abstract and citation database [3]. A total of 401 papers on multimodal interactions were collected from the Scopus database. The purpose of this study was to provide a comprehensive analysis and identification of the recent research trends, including literature types, publication output, national contributions, SCOPUS top categories and journals, top authors, top research areas, and author keywords and analysis. Keyword Plus covers the field of multimodal interactions and its most related areas of research. Therefore, the analysis in this paper makes a significant contribution to researchers interested in multimodal interactions. This paper summarizes trends in multimodal interaction field research and identifies the most relevant research areas to consider in future research. This paper also

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

provides an extensive analysis of related multimodal interaction research fields. Therefore, this study will help researchers identify relevant research areas in multimodal interactions, which have received widespread attention and the gaps that need to be addressed. The structure of this paper is as follows: Section 2 discusses research methodology, Section 3 provides research findings, Section 4 discusses the top 20 most cited papers in the field of multimodal interaction, and Section 5 summarizes this work.

II. METHODOLOGY

Both Web of Science (WoS) and Scopus are the two significant databases used to collect data for scholarly publications [14]. WoS is known as one of the largest and most reliable databases for document retrieval and analysis. However, Scopus' journal coverage appears to be more comprehensive than WoS [15]. To select an appropriate database for research, the main title search of "multimodal interaction" and the topic search of "multimodal interaction" was run on Scopus databases and WoS, respectively. WoS retrieved 348 documents and Scopus database 717 documents. These are the number of papers published from 2012 to 2021, containing the phrase "multimodal interaction" in their titles. Therefore, due to the more significant number of documents in the Scopus database, this data was extracted from the Scopus database on December 31, 2021.

The scope was limited to the publication from 2012 to 2021 to select the latest documents. Therefore, the number of retrieved documents has dropped from 717 to 418. The papers were further restricted to computer science, engineering, mathematics, social sciences, arts and humanities, physics and astronomy, and psychology, with record dropping to 403. Finally, by reading the titles and abstracts, two papers unrelated to multimodal and human-computer interaction (HCI) were excluded. The final data contains 401 documents on multimodal interactions published between 2012 and 2021.

Similar to other bibliometric studies, the data collection procedure of this stud} may have some limitations. The primary dataset consists of publications in the Scopus database that include multimodal interactions in their titles. However, some authors may use synonyms of the phrase instead. However, about 95% of the documents are indexed by the Scopus and WoS databases [16]. WoS is considered one of the largest and most reliable databases for literature retrieval and analysis. However, Scopus' journal coverage appears to be more comprehensive than WoS [17], so the study did not specifically cover literature from the WoS database. Most of the publications in this bibliometric analysis ensure that the literature represents significant research findings in multimodal interaction. Data were analyzed by VOS Viewer software and Bibliometrix-Package4.

VOS Viewer is a software that generates distance-based maps and clusters keywords retrieved from the title and abstract of the study file [16,17]. VOS Viewer software [18] is used to showcase the number of publications per year and visualize the top 20 keywords in multimodal interaction studies.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

Bibliometric software is specifically designed for the quantitative study of scientific indicators and bibliometrics [19,20]. It offers various programs to import bibliographic data from well-known databases such as Scopus and WoS [19,20]. The data acquisition process is shown in Figure 1. The results of the bibliometric analysis are combined with a qualitative analysis of the literature content.

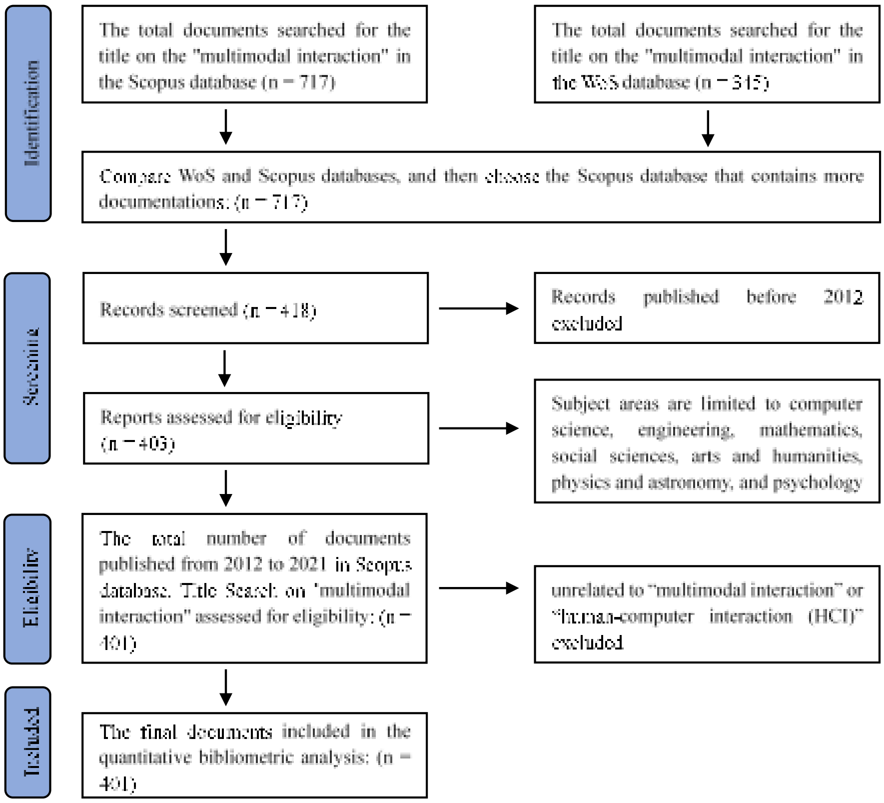

Figure 1. PRISMA flow chart of the multimodal interaction document collection.

Table I shows the primary information about the 401 documents on multimodal interaction in Scopus Database. From 2012 to 2021, there were 247 conference papers, 99 articles, 23 conference review papers, 16 book chapters, nine reviews, and one book on multimodal interaction in the Scopus database. To draw the current sub-topic of multimodal interaction research, we selected the 20 most cited articles for qualitative analysis. The following sections provide quantitative and qualitative analysis of the data.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

Table I. Main information about 401 documents on multimodal interaction in Scopus Database.

Description	Results
Timespan	2012:2021
Sources (Journals. Books. etc]	247
Documents	401
Article	99
Book	1
Book chapter	16
Conference paper	247
Conference review	22
Editorial	7
Review	9
Average citations per document	5.86
Keywords plus (ID)	2100
Author's keywords (DE)	1004
Authors	1193
Authors of single-authored documents	39
Authors of multi-authored documents	1154
Authors per document	2.98
Co-Authors per documents	3.63
Collaboration index	3.4

III. QUANTITATIVE ANALYSIS

3.1 Sources Analysis

Table I shows that there are 247 sources, such as journals, books, and others, in this bibliographic collection set. Figure 2 shows the most relevant sources in the field of multimodal interactions. Each source publishes one or more documents in this analysis set. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) is a top journal in multimodal interaction, with 34 papers from 2012 to 2021. The ACM International Conference Proceedings Series Journal published twelve documents on multimodal interaction, ranking second. Journal on Multimodal User Interfaces has nine. Other journals have published fewer than nine papers in this analysis dataset. The three top journals in Figure 2 are significant for researchers in the multimodal interaction area and can first consider submitting their documents.

Figure 3 shows the number of sources occurrences from 2012 to 2021. The Figure demonstrates the best sources of dynamics in multimodal interaction. Over time, the number of annual publications on top

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

journals has increased. The Journal of Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) is the leader in the most relevant articles. Since 2012, Acm International Conference Proceeding Series and other three journals have been published multimodal interaction-related papers on an upward trend.

Figure 2. The top 20 sources for publishing multimodal interaction documents

3.2 Publication Years Analysis

Figure 4 presents the scientific documents on multimodal interaction areas from 2012 to December 31, 2021. During this period, 401 academic works were published, including 247 conference papers, 99 articles, 22 conference reviews, 16 book chapters, nine review papers, seven editorial, and one book. The number of articles published in 2012 (n=37) is almost the same as in 2021 (n=36). The distribution of multimodal interaction publications is divided into two stages in this process. In the first stage, from 2012 to 2016, the publications continuously rose and peaked at 54 in 2016, even though a little declined in 2015. However, in the second stage, from 2016 to 2021, in 2019, the publications began to fall and reached the bottom 29. The number of published papers dropped from 54 in 2016 to 29 in 2019, but more researchers were devoted to this area in 2020.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

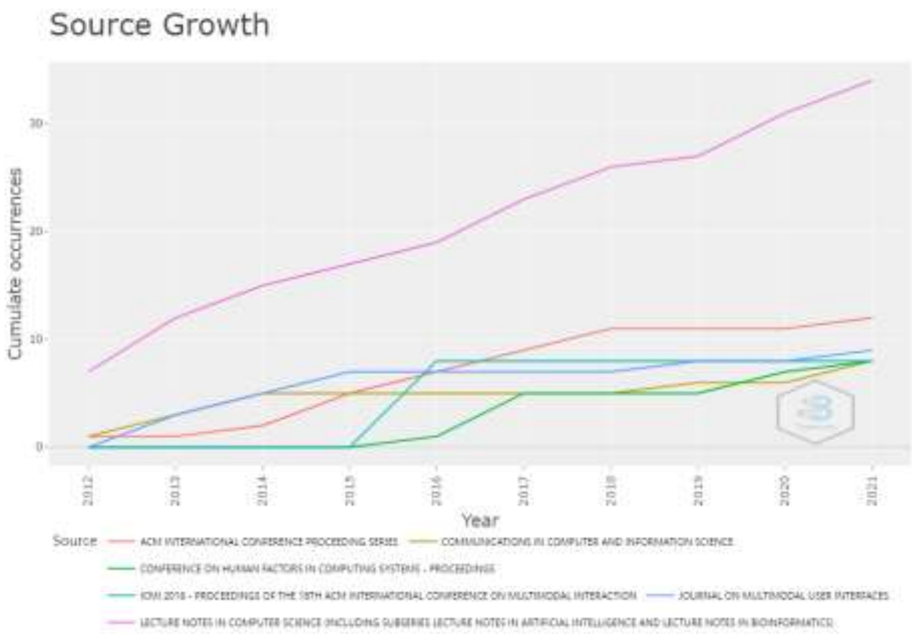

Figure 3. The annual publication of the top 5 Sources in multimodal interaction within 2012 - 2021.

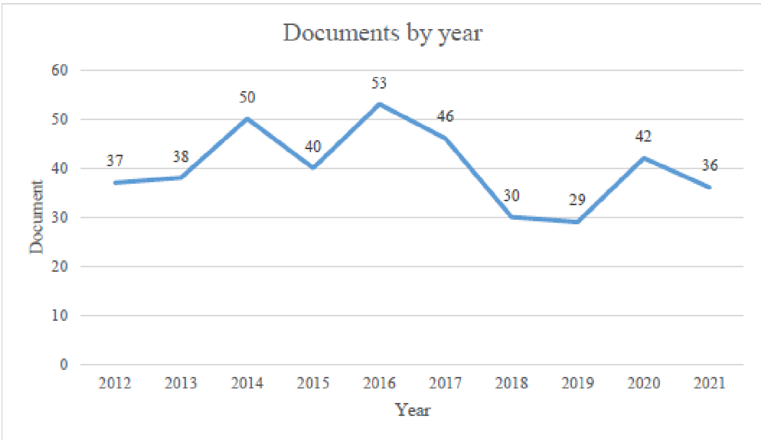

Figure 4 Annual scientific production of multimodal interaction research field from 2012 to 2021.

3.3 Analysis of Authors

As Table I shows, 1193 researchers published their scientific achievements in the multimodal interaction area from 2012 to 2021 in the Scopus database. It means that each document has 3.63 co-authors.

Figures 5 and 6 show the number and span of published articles by the most relevant authors in 2012 -

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

2021. The red line means the author's timeline. The size of the circle is related to the number of documents published over the years. The larger size of the circle represents more published papers during the year. The concentration of circle color is related to the total number of citations per year. Honold. F. is the author with the most documents in the multimodal interaction area, with nine from 2012 to 2017. In the same period. Schussel, F. püblished seven papers, ranking second. Weber, M., Teixeira, a., and Almeida, N. have six publications, respectively, ranking third together. Almeida, N., Teixeira, A., and Fuhrmann, F. have the most protracted timeline among other authors. There are six highly cited papers published from 2012 to 2021. According to Table II and Figure 6, author Srinivasan A. published two highly cited papers in 2018, with 62 total citations (TC), 15.5 total citations per year (TCpY); and published another two highly cited papers in 2020, with 23 total citations (TC), 11.5 total citations per year (TCpY). Stasko J. published two highly cited articles in 2018, with 62 total citations (TC), 15.5 total citations per year (TCpY). The other authors' total citations per year (TCpY) is less than 3.6. So, Srinivasan A. and Stasko J. are the top authors in multimodal interaction research areas. Other researchers in this field must pay attention to their publications.

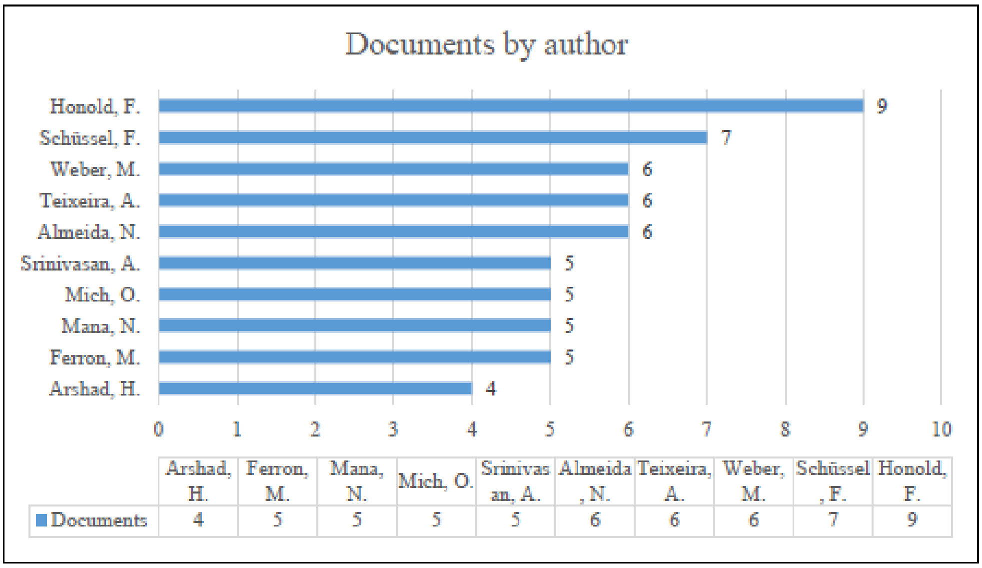

Figure 5 Documents by author

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

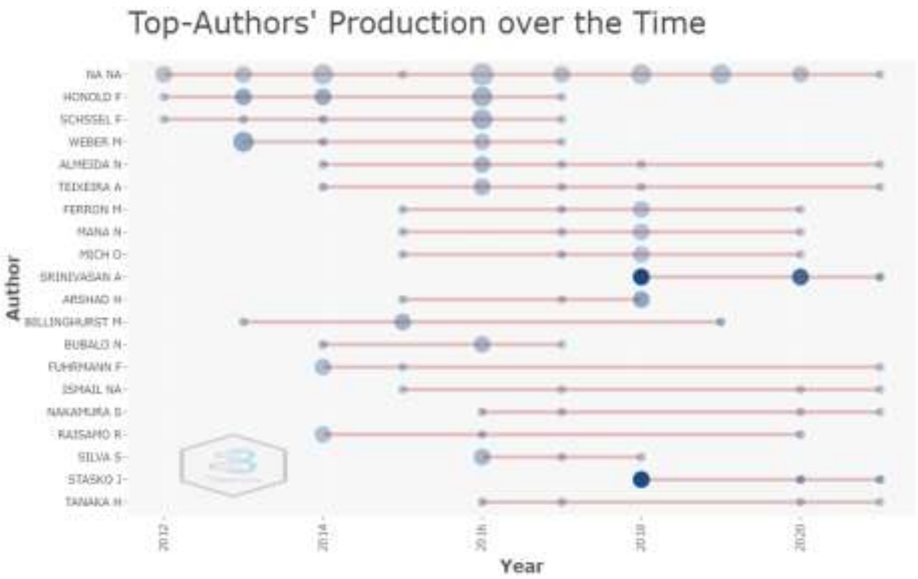

Figure 6 Top-Authors' Production over the Time

Table II. Author Production over Time

Author	Year	Freq	TC	TCpY
SRINIVASAN A	2018	2	62	15.5
STASKO J	2018	2	62	15.5
WEBER M	2013	3	32	3.556
HONOLD F	2013	2	31	3.444
HONOLD F	2014	2	23	2.875
SRINIVASAN A	2020	2	23	11.5
BUBALO N	2014	1	19	2.375
SCHSSEL F	2014	1	19	2.375
WEBER M	2014	1	19	2.375
ALMEIDA N	2014	1	16	2
TEIXEIRA A	2014	1	16	2
ARSHAD H	2018	2	15	3.75
BILLINGHURST M	2015	2	15	2.143
SCHSSEL F	2013	1	14	1.556
BILLINGHURST M	2013	1	12	1.333
HONOLD F	2016	3	11	1.833
SCHSSEL F	2016	3	11	1.833
STASKO J	2020	1	10	5
ALMEIDA N	2016	2	9	1.5

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

Author	Year	Freq	TC	TCpY
SILVA S	2016	2	9	1.5

3.4 Analysis of Document Subject

Figure 7. shows the subject distribution of the multimodal interaction-related work. The highest frequency belonged to “Computer Science” (N=247, 51%), followed by engineering subjects (N=99, 14%), with N=63, 8% appearing in Mathematics subject. The frequency of articles in Social Sciences (N=51, 8%) was greater than that in Arts and Humanities (N=23, 4%), Neuroscience (N=16, 2%), Materials Science (N=15, 2%), Psychology (N=15, 2%), and Physics and Astronomy (N=13, 2%). Publications distributed in other fields N=39, 6%.

The statistical analysis of subject distribution shows that the research on multimodal interaction mainly focuses on Computer Science, Engineering, and Mathematics. The wide distribution of disciplines means that the study of multimodal interaction has received the attention of researchers in various fields, and the research on multimodal interaction is universal.

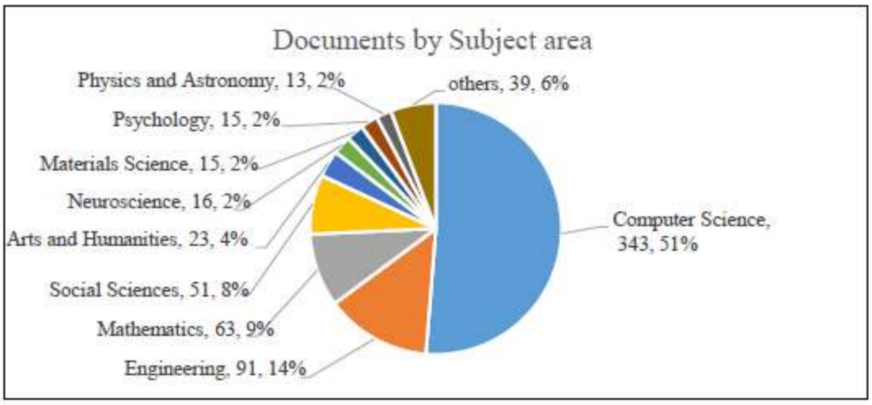

Figure 7. The subject area of the documents.

3.5 Analysis of Documents Types

The results of frequency related to publication types show that the highest frequency belongs to “conference papers” (N=247, 61.6%, followed by article (N=99, 24.7%), conference review (N=22, 5.5%), and book chapter (N=16, 4.0%), as seen in Figure 8.

There are few book chapters (N=16, 4.0%), which means the multimodal interaction is a new research area, and most of the researchers have not compiled the research results into books. The most significant number of publications are conference papers (N=247, 61.6%), which means that the field of multimodal

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

interaction continues attracting the attention of researchers, who have not been able to translate the research results into articles in time. Therefore, although researchers have studied the field of multimodal interaction for a long time, it has always been a research hotspot.

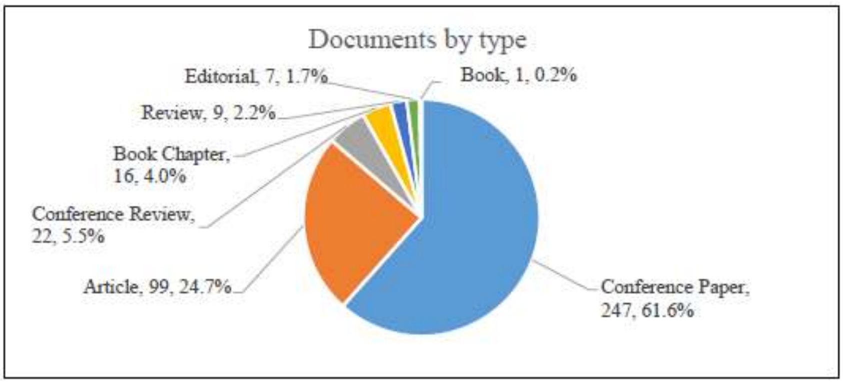

Figure 8. Type of document

3.6 Analysis of Topics

Keyword networks are used to analyze the main topics of multimodal interaction publications. Keyword networks represent co-occurrence between bibliographic datasets. It is possible to highlight various topics by clustering keyword networks. Each keyword belongs to a topic. Therefore, the thematic map represents a specific map for each topic. Each circle shows a keyword network cluster, and the cluster name is the word with the highest existence rate. Figure 9 shows the thematic map in multimodal interaction. Therefore, "Multimodal interaction," "Multimodal," "Multimodal interface," "Human-computer interaction," and "Automation" are the most relevant subject indicators.

The first quadrant (upper right) is the motion theme, indicating importance and perfection. The second quadrant (upper left) represents highly developed and isolated topics that have developed well but are not critical to the current field. The third quadrant lower-left corner) means emerging or declining themes. There is no good development. It may have just occurred or maybe about to disappear. The fourth quadrant (lower-right corner) represents basic and transversal themes, very important to the field, but not well developed, generally, refers to basic concepts. The size of the circle is related to the emergence of clustering words, and its position depends on the centrality and intensity of clustering. The centrality and intensity show the importance and improvement in multimodal interaction.

On the other hand, keywords such as "multimodal interactions," "interactive computer system," "and user interfaces" are in one cluster, while "multimodal," "virtual reality," and "speech recognition" are in another set. The representatives of these two groups are called the basic theme and the horizontal theme. Among them, the keyword topic of "multimodal interactions" is the most essential and core topic.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

"Multimodal interactions," "interactive computer systems," and "user interface" are the main keyword themes, and they are more likely to be used in future research on multimodal interaction.

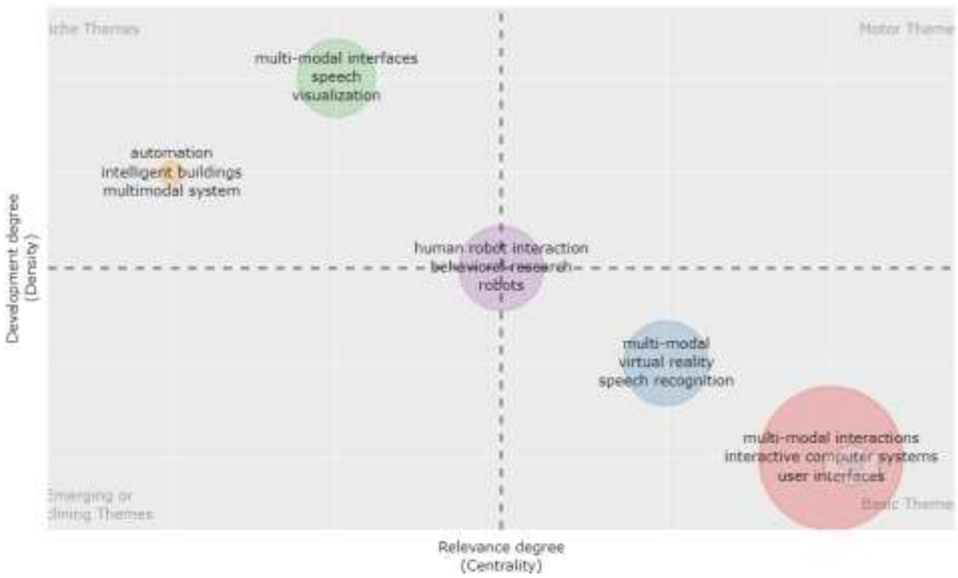

Figure 9. Thematic Map of publication in multimodal interaction research field within 2012 to 2021

3.7 Geographical Distribution Analysis of Publications

The geographical distribution of publications has become an exciting indicator of the productivity of a country, region, or institution. The analysis results related to the geographical distribution of contributions in this field (Figure 10) show that the most significant number of publications related to multimodal interaction is the United States (64), followed by Germany (52). The contribution of Japan was 34. While Austria, Australia, Denmark, Finland, Greece, the Netherlands, and other 40 countries have less than ten publications related to multimodal interaction. Over 50 countries contributed to the 401 publications. This regional distribution of multimodal interaction-related research is indicative of general publication trends in most fields and disciplines.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

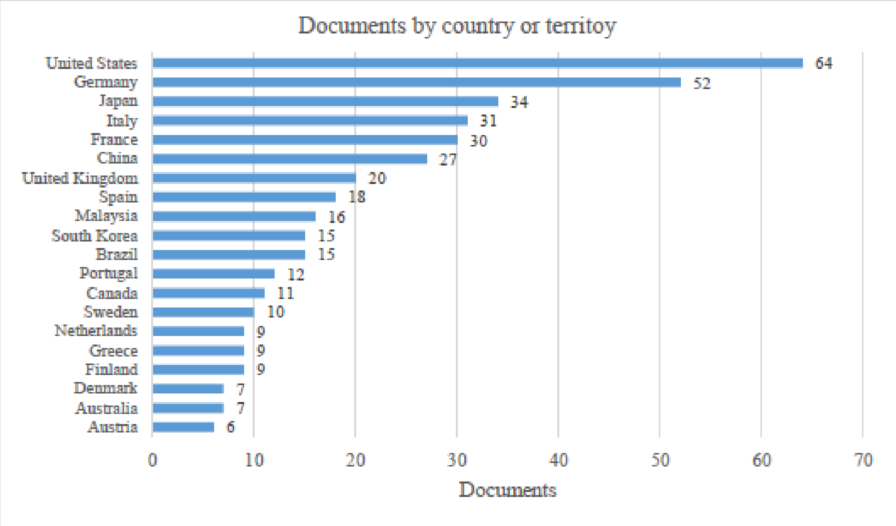

Figure 10. Geographical distribution of publication for Scopus databases from 2012 to 2021

3.8 Analysis of the Top Twenty Publications

Seventy-seven authors contributed 20 of the most cited articles from 2012 to 2021 (Table III). The most frequently cited paper was published in 2014 in the Scopus database, receiving 213 citations; eight out of the twenty top publications were published in computer science-related journals; and seven of which are conference proceedings.

The top ten most-cited publications discussed the application of multimodal interaction and the review of multimodal interaction research. The geographical distribution of multimodal interaction-related work indicates that discussion on multimodal interaction is universal.

Table III. Top twenty publications based on the Times citation per year for Scopus databases from 2012 to 2021

No.	Authors	Title	Year	Source title	TC	TCpY
1	Turk, M.	Multimodal interaction: A review [21]	2014	Pattern Recognition Letters	213	30
2	Fernandez, R. A. S., Sanchez-Lopez, J. L., Sampedro, C., Bavle, H., Molina, M., & Campoy, P.	Natural user interfaces for human-drone multimodal interaction [22]	2016	International UAV Systems Conference 2016, ICAUS 2016	88	18
3	Srinivasan A., Stasko J.	Orko: Facilitating multimodal interaction for visual exploration and analysis of networks [23]	2018	IEEE Transactions on Visualization and Computer Graphics	52	17

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

No.	Authors	Title	Year	Source title	TC	TCpY
4	Buehlmann C., Mangan M., Graham P.	Multimodal interactions in insect navigation [24]	2020	Animal Cognition	16	16
5	Papanastasiou, S., Kousi, N., Karagiannis, P., Gkournelos, C., Papavasileiou, A., Dimoulas, K., ... & Makris, S.	Towards seamless human robot collaboration: integrating multimodal interaction [25]	2019	International Journal of Advanced Manufacturing Technology	28	14
6	Srinivasan, A., Lee, B., Henry Riche, N., Drucker, S. M., & Hinckley, K.	InChorus: Designing consistent multimodal interactions for data visualization on tablet devices [26]	2020	Conference on Human Factors in Computing Systems - Proceedings	13	13
7	Deppermann, A.	Multimodal interaction from a conversation analytic perspective [27]	2013	Journal of Pragmatics	96	12
8	Martinez-Maldonado, R., Kay, J., Buckingham Shum, S., & Yacef, K.	Collocated collaboration analytics: Principles and dilemmas for mining multimodal interaction data [28]	2019	Human-Computer Interaction	20	10
9	Saktheeswaran A., Srinivasan A., Stasko J.	Touch? Speech? or Touch and Speech? Investigating Multimodal Interaction for Visual Network Exploration and Analysis [29]	2020	IEEE Transactions on Visualization and Computer Graphics	10	10
10	Hampel R., Stickler U.	The use of videoconferencing to support multimodal interaction in an online language classroom [30]	2012	ReCALL	89	10
11	Liu, C., Lin, Z., Shen, X., Yang, J., Lu, X., & Yuille, A.KIM	Recurrent Multimodal Interaction for Referring Image Segmentation [31]	2017	Proceedings of the IEEE International Conference on Computer Vision	39	10
12	Morris M.R.	Web on the wall: insights from a multimodal interaction elicitation study [32]	2012	ITS 2012 - Proceedings of the ACM Conference on Interactive Tabletops and Surfaces	86	10
13	Kim, S. S., Gomez-Ramirez, M., Thakur, P. H., & Hsiao, S. S.	Multimodal interactions between proprioceptive and cutaneous signals in primary somatosensory cortex [33]	2015	Neuron	53	9
14	Kerzel, M., Strahl, E., Magg, S., Navarro-Guerrero, N., Heinrich, S., & Wermter, S.	NICO—Neuro-inspired companion: A developmental humanoid robot platform for multimodal interaction [34]	2017	RO-MAN 2017 - 26th IEEE International Symposium on Robot and Human Interactive Communication	34	9
15	Gerdes, A., Wieser, M. J., & Alpers, G. W.	Emotional pictures and sounds: a review of multimodal interactions of emotion cues in multiple domains [35]	2014	Frontiers in Psychology	53	8
16	Cacace, J., Finzi, A., Lippiello, V., Furci, M., Mimmo, N., & Marconi, L.	A control architecture for multiple drones operated via multimodal interaction in search & rescue mission [36]	2016	SSRR 2016 - International Symposium on Safety, Security and Rescue Robotics	37	7

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

No.	Authors	Title	Year	Source title	TC	TCpY
17	Pfleging, B., Schneegass, S., & Schmidt, A.	Multimodal interaction in the car: combining speech and gestures on the steering wheel [37]	2012	Automotive UI 2012 - 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, In-cooperation with ACM SIGCHI - Proceedings	61	7
18	Cohn, N.	A multimodal parallel architecture: A cognitive framework for multimodal interactions [38]	2016	Cognition	31	6
19	Gürkök, H., & Nijholt, A.	Brain-computer interfaces for multimodal interaction: A survey and principles [39]	2012	International Journal of Human-computer Interaction	55	6
20	Makri, E., Spiliotopoulos, D., Vassilakis, C., & Margaris, D.	Human behavior in multimodal interaction: main effects of civic action and interpersonal and problem-solving skills [40]	2020	Journal of Ambient Intelligence and Humanized Computing	5	5

3.9 Visualize Topic Analysis across Databases

Figure 11 summarizes the frequency of different topics and keywords in the 401 Scopus topic categories. In general, the proximity of two subject types reflects the strength of their relevance, while the size of one subject category indicates the number of publications on the subject. Colors range from blue to red and indicate the publication time of the subject category, blue for earlier publications, and red for newer publications. As figure 11 shows, the earlier published work on multimodal interactions in the Scopus database tends to be closely related to issues such as "multimodality," "multimodal," "interaction," "multimodal interface," and "design.". Among recent publications, the multimodal interaction research publications in Scopus are related to “Behavioral Research,” “Augment Reality,” “Visualization,” “Gesture Recognition,” and “Human engineering.” Publications after 2016 focus more on “visualization,” “augmented reality.”

The map of trends in Scopus in Figure 11 shows that work on multimodal interaction is moving towards research focusing on specific subject areas and specific applications.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

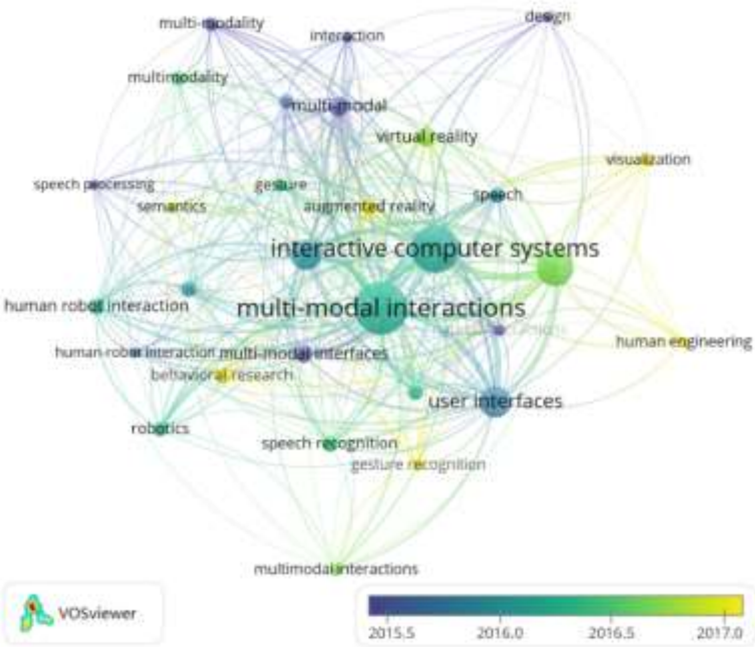

Figure 11 Map of articles during 2012 –2021, bibliographic coupling, multimodal interaction in Scopus

IV. DISCUSSION

Bibliometrics has the advantage of solid objectivity and can exclude the influence of subjective factors on research results. From this perspective, bibliometrics is superior to peer review. However, bibliometrics also has some unavoidable disadvantages. To avoid these disadvantages, qualitative knowledge of the scholars involved and their active branches should combine with objective data from bibliometrics [41].

The start-point and end-point of the timeline represent the rise and fall of research topics, and can also reflect research hotspots in different periods. The length of the timeline represents the duration of popular research topics. Based on this, the trends of the top 36 research topics from 2012 to 2021 are derived. The size of the circle represents how often the topic appears. As shown in Figure 12, the research hotspots in different periods have changed differently. For example, the hotspot in 2012 was “Artificial intelligence,” and it lasted until 2016. In 2014-2018, "multimodal interactions," "interactives computer systems," and "human-computer interaction" are the main topics, and "multimodal interactions" with the highest frequency. The topics of "human," "mammals," and "behavioral research" have the most extended timeline. Interestingly, "Multimodal Interactions," a research hotspot in 2016, again became the research hotspot in 2021. The research topic trends are consistent with session 3.6 Analysis of Topics.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

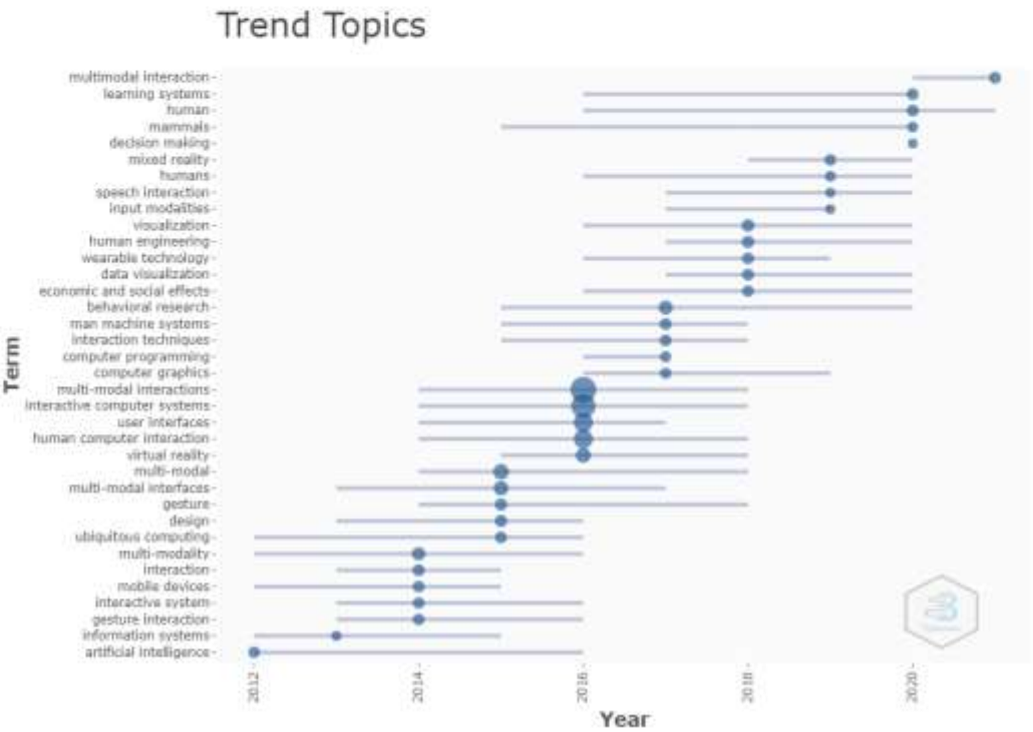

Figure 12. Map of trend Topics, during 2012 –2021, in the multimodal interaction research area in Scopus database

It is necessary to discuss the top 20 pieces of literature to explore the latest trends and research hotspots related to multimodal interactions, including augmented reality, human-robot, and behavioral, to keep up with the latest developments in academic research.

Humper et al. studied how multimodal online video conferencing can be applied to language teaching and how teachers and students can adapt to multimodal online situations [30]. Morris et al. interviewed 25 participants to learn about scenarios where they wanted to use a web browser on the TV in the living room [32]. Pflegging et al. Compared the effects of multimodal interaction (combination of gesture and sound) and physical buttons on driving performance. The findings showed that using voice and gesture was faster than using physical buttons; operator performance was similar, but interaction efficiency was low; when using sound and gesture, the requirements for vision were lower than those for using physical interaction [37]. Cohn et al. integrated multimodal interaction into existing language and cognitive frameworks and described the interactions between different complexities in the domains of language, body, and graphics [38]. The resulting theoretical model can be beneficial to the study of corpus analysis, experimentation, and multimodal educational benefits. Makri et al. have extensively studied multimodal interactions (virtual agents, verbal conversations, progressive visual communication) in metacognitive skills training [40]. The results show a significant positive correlation between users and system evaluation, self-efficacy, self-

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

regulation, personal change preparation, goal-oriented mastery, interpersonal communication, and problem-solving ability after interacting with the dialogue system.

Robot technology is considered one of the high-tech technologies that have an essential impact on emerging industries in the future [28]. Robotics research has expanded from traditional industrial fields to new technologies such as medical services, education and entertainment, exploration and measurement, biological engineering, and disaster relief and has developed rapidly. Fernandez et al. researched and implemented a graphical user interface (GUI) and several NUI methods and computer vision techniques in an aerial robot software framework named Aerostack [22]. Cacace et al. proposed an architecture suitable for controlling multiple UAVs deployed in search and rescue missions [36]. Kerzel et al. developed and launched a humanoid development robot Nico (neurologically inspired Companion), filling the gap between the necessary perception and interaction capabilities and flexible design [34]. Papanastasiou et al. discuss the challenges of human operators and industrial robots collaborating in assembly operations, focusing on safety and simplified interactions [25].

Visualization mainly refers to the technically visual interpretation of data, a relatively mature research field. Martinez Maldonado et al. introduced and proposed a set of principles for mining juxtaposed collaborative data [28]. Liu et al. proposed a convolutional multimodal LSTM to encode sequential interactions between a single word, visual information, and spatial information [31]. The results show that the model is superior to the baseline model in the benchmark data set. Srinivasan et al. Created architecture and a visualization system prototype called Orko to promote natural language and direct operational input [23]. Srinivasan et al. based on a set of core concepts (including operations, parameters, objectives, instruments) and combing with standard interface elements (such as axes, markers), developed an interactive system suitable for different types of visualization [26]. Saktheeswaran et al. conducted qualitative user research in network visualization tools and compared a single-mode interface based on voice and touch with a multi-mode interface that combines both. The findings approved that multimodal input was more favored by participants than single-mode input [29]. Gürkök et al. investigated the most advanced real-time multi-mode BCI application. The results show that multimodal use of BCI can help to improve error handling ability, task performance, user experience and expand the scope of users [39]. Gerdes et al. outline research on emotional sounds and interaction with pictures in complex scenes [35].

ACI (Animal-Computer Interaction) is a new, fascinating, and potential interdisciplinary field that has increasingly attracted the attention of researchers. Srinivasan et al. introduced the current knowledge of using multimodal cues in insect orientation and navigation. Their results found that multimodality can improve robustness, accuracy, and overall foraging success rate [24]. Kim et al. used a mechanical stimulator that can reliably position the animal's hand in different postures and present tactile stimuli with extremely high precision. Their experimental data provide evidence for other mechanisms of multimodal processing in somatosensory systems [41].

Multimodal research based on voice, touch, vision, and gesture has been on the rise in recent years with the development of non-desktop computing generated by mobile devices and reasonably priced sensors [42]. A sub-topic within HCI is Multimodal Interaction (MMI), which combines multiple modes to provide a powerful, flexible, adaptable, and natural interface [24]. Multimodality is an essential direction towards general artificial intelligence [43]. Multimodal fusion is the future development trend [44]. The future human-computer interaction system will be more intelligent, convenient, rational, and emotional.

V. CONCLUSION AND FUTURE WORK

Mathematical and statistical are usually used for bibliometric analysis and evaluation of academic achievements. This work focuses on bibliometric analysis of multimodal interaction research literature and Scopus data sources from 2012 to 2021. The findings show significant progress has been made in the basic research and application of multimodal interaction in recent years. Although, the number of documents published in multimodal interaction research has decreased from 54 in 2016 to 29 in 2019. In 2020, more researchers focused on this field. Researchers from the USA, Germany, Italy, Japan, China, France, and the United Kingdom contributed the most. The USA and Germany ranked first and second in paper production, respectively. Computer Science Lectures (Including Artificial Intelligence Sub-Series Lectures and Bioinformatics Lectures), Acm International Conference Series Journals, and Multimodal User Interface Journals have been identified as the most representative journal in the multimodal interaction research area. The main topics in our bibliography are representative keywords, such as "multimodal interaction," "interactive computer system," and "user interface." In recent years, keywords such as multimodal interaction and multimodality are increasing their visibility. It is expected that this work will provide good prospects for future research and help researchers discover potential opportunities in multimodal interaction.

It can be predicted that there will be many researchers dedicated to the above hot research topics and will produce many research achievements.

REFERENCES

[1] Anonymous. (December 2021). What is multimodal interaction?. Retrieved from https://blog.csdn.net/qq_14942529/article/details/113913799

[2] Liu Yunfeng. (December 2021). Multimodal digital human, human-computer interaction in the 5G era (zhihu.com) Retrieved from https://zhuanlan.zhihu.com/p/ 213801908.

[3] Elsevier. (December 2021). How Scopus Works. Retrieved from https://www.elsevier.com/solutions/scopus/how-scopus-works/content?dgcid=RN_ AGCM_Sourced_300005030.

[4] Waltman L. (2016). A review of the literature on citation impact indicators. Journal of informetrics, 10(2), 365-391.

[5] Kalantari A, Kamsin A, Kamaruddin H S, Ebrahim N A, Gani A, Ebrahimi A, Shamshirband S. (2017). A bibliometric approach to tracking big data research trends. Journal of Big Data, 4(1), 1-18.

[6] Repanovici A. (2010, August). Measuring the visibility of the university's scientific production using google scholar, Publish or Perish Software, and Scientometrics. In World Library and Information Congress: 76th IFLA General Conference and Assembly. Retrieved December (Vol. 19, p. 2010). Gothenburg, Sweden: Science and Technology Libraries.
[7] Zitt M, Ramanana-Rahary S, Bassecoulard E. (2005). Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalization. Scientometrics, 63(2), 373-401.
[8] Li L L, et al. (2009).Global stem cell research trend: Bibliometric analysis as a tool for mapping of trends from 1991 to 2006. Scientometrics, 80(1): 39-58.
[9] Ale Ebrahim N, Salehi H, Embi M A, Habibi F, Gholizadeh H, Motahar S M. (2014). Visibility and citation impact. International Education Studies, 7(4), 120-125.
[10] Budd J M. (1988). A Bibliometric Analysis of Higher-Education Literature. Research in Higher Education, 28(2): 180-190.
[11] Cañas-Guerrero I, Mazarrón F R, Calleja-Perucho C, Pou-Merina A. (2014). Bibliometric analysis in the international context of the "Construction & Building Technology" category from the Web of Science database. Construction and Building Materials, 53, 13-25.
[12] Cañas-Guerrero I, Mazarrón F R, Pou-Merina A, Calleja-Perucho C, Díaz-Rubio G. (2013). Bibliometric analysis of research activity in the "Agronomy" category from the Web of Science, 1997–2011. European Journal of Agronomy, 50, 19-28.
[13] Ingwersen P. (2000). The international visibility and citation impact of Scandinavian research articles in selected Social Science fields: The decay of a myth. Scientometrics, 49(1): 39-61.
[14] Aghaei Chadegani A, Salehi H, Yunus M, Farhadi H, Fooladi M, Farhadi M, Ale Ebrahim N. (2013). A comparison between two main academic literature collections: Web of Science and Scopus databases. Asian social science, 9(5), 18-26.
[15] Agarwal A, Durairajanayagam D, Tatagari S, Esteves S C, Harlev A, Henkel R, Bashiri A. (2016). Bibliometrics: tracking research impact by selecting the appropriate metrics. Asian journal of andrology, 18(2), 296.
[16] Khalil G M, Crawford C A G. (2015). A bibliometric analysis of US-based research on the behavioral risk factor surveillance system. American journal of preventive medicine, 48(1), 50-57.
[17] Norfarah N, Ali S M, Siti-Nabiha A K, Ebrahim N A. (2019). Firms' sustainable practice research in developing countries: mapping the cited literature by bibliometric analysis approach. International Journal of Sustainable Strategic Management, 7(1-2), 5-26.
[18] Van Eck N J, Waltman L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics, 84(2), 523-538.
[19] Aria M, Cuccurullo C. (2017). bibliometrics: An R-tool for comprehensive science mapping analysis. Journal of informetrics, 11(4), 959-975.
[20] Moher D, Liberati A, Tetzlaff J, Altman D G, Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS medicine, 6(7), e1000097.
[21] Turk M. (2014). Multimodal interaction: A review. Pattern recognition letters, 36, 189-195.
[22] Fernandez R A S, Sanchez-Lopez J L, Sampedro C, Bavle H, Molina M, Campoy P. (2016, June). Natural user interfaces for human-drone multimodal interaction. In 2016 International Conference on Unmanned Aircraft Systems (ICUAS) (pp. 1013-1022). IEEE.
[23] Srinivasan A, Stasko J. (2017). Orko: Facilitating multimodal interaction for visual exploration and analysis of networks. IEEE transactions on visualization and computer graphics, 24(1), 511-521.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

[24] Buehlmann C, Mangan M, Graham P. (2020). Multimodal interactions in insect navigation. Animal cognition, 1-13.
[25] Papanastasiou S, Kousi N, Karagiannis P, Gkournelos C, Papavasileiou A, Dimoulas K, Makris S. (2019). Towards seamless human-robot collaboration: integrating multimodal interaction. The International Journal of Advanced Manufacturing Technology, 105(9), 3881-3897.
[26] Srinivasan A, Lee B, Henry Riche N, Drucker S M, Hinckley K. (2020, April). InChorus: Designing consistent multimodal interactions for data visualization on tablet devices. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-13).
[27] Deppermann A. (2013). Multimodal interaction from a conversation analytic perspective. Journal of Pragmatics, 46(1): 1-7.
[28] Martinez-Maldonado R, Kay J, Buckingham Shum S, Yacef K. (2019). Collocated collaboration analytics: Principles and dilemmas for mining multimodal interaction data. Human-Computer Interaction, 34(1), 1-50.
[29] Saktheeswaran A, Srinivasan A, Stasko J. (2020). Touch? Speech? or Touch and Speech? Investigating multimodal interaction for visual network exploration and analysis. IEEE transactions on visualization and computer graphics, 26(6), 2168-2179.
[30] Hampel R, Stickler U. (2012). The use of videoconferencing to support multimodal interaction in an online language classroom. ReCALL, 24(2), 116-137.
[31] Liu C, Lin Z, Shen X, Yang J, Lu X, Yuille A. (2017). Recurrent multimodal interaction for referring to image segmentation. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1271-1280).
[32] Morris M R. (2012, November). Web on the wall: insights from a multimodal interaction elicitation study. In Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces (pp. 95-104).
[33] Kim S S, Gomez-Ramirez M, Thakur P H, Hsiao S S. (2015). Multimodal interactions between proprioceptive and cutaneous signals in primary somatosensory cortex. Neuron, 86(2), 555-566.
[34] Kerzel M, Strahl E, Magg S, Navarro-Guerrero N, Heinrich S, Wermter S. (2017, August). NICO—Neuro-inspired companion: A developmental humanoid robot platform for multimodal interaction. In 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) (pp. 113-120). IEEE.
[35] Gerdes A, Wieser M J, Alpers G W. (2014). Emotional pictures and sounds: a review of multimodal interactions of emotion cues in multiple domains. Frontiers in Psychology, 5, 1351.
[36] Cacace J, Finzi A, Lippiello V, Furci M, Mimmo N, Marconi L. (2016, October). A control architecture for multiple drones operated via multimodal interaction in search & rescue missions. In 2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR) (pp. 233-239). IEEE.
[37] Pfleging B, Schneegass S, Schmidt A. (2012, October). Multimodal interaction in the car: combining speech and gestures on the steering wheel. In Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 155-162).
[38] Cohn N. (2016). A multimodal parallel architecture: A cognitive framework for multimodal interactions. Cognition, 146, 304-323.
[39] Gürkök H, Nijholt A. (2012). Brain-computer interfaces for multimodal interaction: A survey and principles. International Journal of Human-Computer Interaction, 28(5), 292-307.
[40] Makri E, Spiliotopoulos D, Vassilakis C, Margaris D. (2020). Human behavior in multimodal interaction: main effects of civic action and interpersonal and problem-solving skills. Journal of Ambient Intelligence and Humanized Computing, 11(12), 5991-6006.
[41] Moed H F, Luwel M, Nederhof A J. (2002). Towards research performance in the humanities. Library Trends, 50(3), 498-520.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
Sept-Oct 2021 Page No. 1615-1636
Article History: Received: 10 August 2021, Revised: 25 August 2021, Accepted: 05 September 2021, Publication: 31 October 2021

[42] Xu Y S. (2010). Intelligent robots lead the development of high and new technology. Enterprise Science Association, (9), 28-31.
[43] Yu D. (December 2021). Multimodality is an important direction towards general artificial intelligence. Retrieved from https://cloud.tencent.com/developer/ article/1505461.
[44] Tan H, Sun J, Wenjia W, Zhu C. (2021). User Experience & Usability of Driving: A Bibliometric Analysis of 2000-2019. International Journal of Human-Computer Interaction, 37(4), 297-307.