The Analysis of Research Trends in Biometrics in South Korea

Jinhyo Jung, Hyunjung Yang, and Chang-Moo Lee

***Abstract*—Biometrics has been under the spotlight in academia for long. A great number of academic papers have been published over the past few decades. However, most of them have focused too much on technical issues. There has been only limited research conducted from the comprehensive and overall viewpoint. It is overriding now to categorize and survey the hitherto published papers regarding biometrics to extend the scope of research and facilitate efficient studies. This article aims to shed light on the research streams on biometrics with emphasis on five categories: fingerprint recognition, iris recognition, face recognition, speaker—or speech—recognition, and vein recognition. The author collected 956 papers spanning from 1980 to 2015 through RISS—the largest academic database in Korea, which has a pool of more than four million articles. Based on social network analysis as a main methodology, this survey used analytic and statistical tools— Krwords, Ktitle, R, and Netminer— to analyze and categorize research trends on the selected five categories. This article will contribute to understanding the status quo of biometrics objectively and helping researchers design their own studies about biometrics.**

***Index Terms*—Big data, biometrics, South Korea, social network analysis, research trends.**

I. Introduction

A large amount of research on biometrics has been conducted in recent years, particularly in South Korea. Approximately 1,000 academic papers about biometrics which focused mainly on technologies, algorithms, regulations, standardizations, and applications have been published [1]-[4]. These studies have resulted in improving biometrics both in technologies and markets [5]. However, there has been no significant output in the analysis of the research trends in biometrics. In order to develop biometrics as a discipline, a process to organize and summarize research results is necessary [6]. It's an essential part to describe the overall structure of biometrics to make further studies much more efficient and effective. Thus, the primary objective of this study is to analyze the Korean research trends in biometrics with reference to fingerprint recognition, iris recognition, face recognition, speaker—or speech—recognition, and vein recognition through semantic network analysis. This approach provides the scholars who may struggle to make sense of biometrics field or prepare for their own research with insights into the foundation of biometrics.

Existing research trends have mainly employed qualitative research methods—such as literature research, expert evaluation, and Delphi method—which require a lot of time and money to deal with a huge amount of data. Furthermore, those methods are likely to be influenced by the biased and subjective viewpoints of authors. To move beyond these weaknesses, many researchers have instead used quantitative research methods like data mining (Hyun-jung Kim, Nam-ok Jo & Kyung-shik Shin, 2014). In this paper, research trend analysis is carried out by utilizing the network analysis based on data mining.

Manuscript received September 17, 2017; revised May 1, 2018. This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2017-2014-0-00636) supervised by the IITP(Institute for Information & communications Technology Promotion)

Jinhyo Jung is with Department of Security Convergence, Chung-Ang University, Seoul, South Korea (e-mail: sosilion@gmail.com).

Hyunjung Yang is with Department of Security Convergence, Chung-Ang University, Seoul, South Korea (e-mail: hj20126@naver.com).

Chang-Moo Lee is with Department of Industrial Security, Chung-Ang University, Seoul, South Korea (e-mail: jbalanced@gmail.com).

II. Literature Review

A. Research Trends in Biometrics

Biometrics refers to the automated recognition of individuals based on their biological and behavioral characteristics such as fingerprint, face, iris, voice, vein, and a way of walking [7]-[9]. Because the biological information is considered inherent to an individual, there is little risk of being stolen or copied by others. For these strengths, it is used as a safe and effective method of identifying individuals [10], [11]

The growth in studies on biometrics has been dramatic, resulting in a sharp increase in the number of academic papers since the 1990s. Of all the academic papers regarding biometrics, however, the majority of the papers have paid too much attention to technologies and algorithms. Relatively, very few studies have been devoted to research trends in biometrics. Kim, K.-S. and Kim, D. U.[9] briefly introduced the strong points and weaknesses of fingerprint recognition, face recognition, long sentence recognition, iris recognition, retina recognition, speech recognition, and handwriting recognition; and analyzed the current status of markets, both domestically and globally. Yun, Sung-Min [4] compared and analyzed the access control systems with a focus on fingerprint recognition and iris recognition—in the Netherlands, the United States, Canada, Australia, the United Kingdom, Germany, Singapore, Hong Kong, and Portugal. Hong Seung June [3] analyzed the current state of the biometric markets in Korea. This study focused mostly on face recognition, fingerprint recognition, iris recognition. Although some of the papers analyzed several types of biometrics, existing research trends dealt mainly with

doi: 10.18178/ijssh.2018.8.5.952

technology, policy, market, and industry. As the references indicate, the number of such papers is insufficient, and those studies are no more than summaries without specific research methods. Now, a systematic study based on an appropriate method that provides a full understanding and describes the flow of the development is needed.

B. Semantic Network Analysis

Semantic network analysis is a method of analyzing the relationship between words in texts and of grasping the overall structure. The relationship is visualized as a network which shows the connection between words by using lines—called a "link". The relationship between words is analyzed so that the meaning and importance of the words in the whole network can be grasped in the structural relation [12]-[13].

Most of the research trends that take advantage of network analysis have employed titles or keywords of papers: Lee, Hyang-ee [6] gathered 5,674 domestic dissertations regarding Korean history and analyzed research trends about this subject by using papers' titles as data, and Jane Cho [12] as well used titles of the 1,752 papers about library and information science.

By using network analysis, recent analyses of research trends have obtained more objective information than existing research trend analysis. However, most network analyses have depended too much on titles and keywords, which represent limited information. To solve these problems, abstracts, which reveal much more information, should be exploited as data for the network analysis.

III. Methodology

A. Data

The data, used in this paper, for the semantic network analysis are abstracts of the academic papers from 1980 to 2015. For gathering those data, the author took advantage of an academic searchable data base—RISS—using the following search words: fingerprint recognition [authentication], Iris recognition [authentication], face recognition [authentication], speech recognition [authentication], speaker recognition [authentication], and vein recognition [authentication]. After this process, 956 papers were obtained from 82 academic journals: 125 papers for fingerprint recognition, 39 for Iris recognition, 305 for face recognition, 482 for speaker recognition, and 5 for vein recognition.

B. Period Classification

Each period of the research trends was classified as follows: the fingerprint recognition—from 1990 to 2001 (15 papers), 2002 to 2008 (65 papers), and 2009 to 2015(57 papers); the Iris recognition—from 2002 to 2008 (17 papers) and 2009 to 2015 (20 papers); the face recognition—from 1990 to 2001 (24 papers), 2002 to 2008 (130 papers), and 2009 to 2015 (151 papers); and the speaker recognition including speech recognition—from 1984 to 1989 (16 papers), 1990 to 2001 (191 papers), 2002 to 2008 (145 papers), and 2009 to 2015 (130 papers). With the number of papers varying, depending on the divided periods, the length of each classified period differed. Exceptionally, the vein recognition is excluded from this network analysis, because of lack of the published papers—only five papers.

C. Cleaning Process

After collecting abstracts, cleaning process should be carried out in which stop words—articles, prepositions, conjunctions, adverbs, and other words irrelevant to this research—must be eliminated. For this process the software Krwords and R, which are computer-based content analysis softwares, were employed. Krwords reads texts and lists all of the words written in the texts, by counting the frequencies of the words. From these results, R removes stop words, and Krwords makes the lists of the occurring words once again to choose common keywords for the next process.

D. Common Keywords

Next step was to choose common keywords, which are considered closely related to four types of biometrics, based on frequency of occurring words on the list delivered by Krwords. The criterion for frequency depends on the number of published papers in each limited period. Words mentioned more than five times on the list were classified as a common keyword, in case less than 100 papers were published in that period—i.e. fingerprint recognition from 1990 to 2001 (15 papers), 2002 to 2008 (65 papers), 2009 to 2015 (57 papers); Iris recognition from 2002 to 2008 (17 papers), 2009 to 2015 (20 papers); face recognition from 1990 to 2001 (24 papers); speaker—or speech—recognition from 1994 to 1989 (16 papers). On the other hand, words mentioned more than ten times in over 100 papers were classified as common keywords—i.e. face recognition from 2002 to 2008 (130 papers), 2009 to 2015 (151 papers); speaker—or speech—recognition from 1990 to 2001 (191 papers), 2002 to 2008 (145 papers), 2009 to 2015 (130 papers).

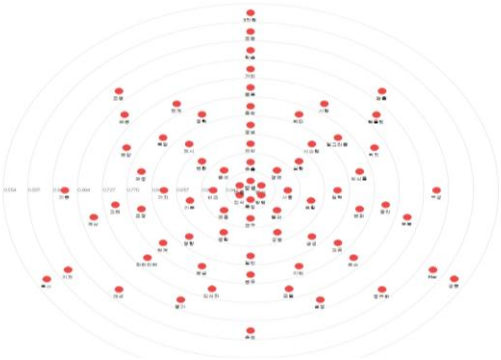

Fig. 1. An example of network visualization of face recognition from 1990 to 2001

E. Network Analysis

As an input data for network analysis, co-occurrence symmetric matrix made up of common keywords selected above were created by a software, Ktitle. This software creates a words-by-words matrix with each cell containing the frequency of co-occurrence of any two words. The classified periods had their own matrixes. Drawing on these matrixes, the author of this article examined closeness centralities—the important indicators for the research trends—by employing

Netminer. This program analyzes centrality and presents network visualization. Analyzed closeness centralities were visualized in the form of cocentric network map, which offers a series of threshold value of the closeness centrality.

TABLE I: AN EXAMPLE OF THE SEMANTIC RELATIONSHIP BETWEEN MAIN KEYWORDS AND COMMON KEYWORDS OF FACE RECOGNITION FROM 1990 TO 2001

	Main Keywords	Closeness Centrality	The Three Top Common Keywords		
			Case Frequency		
1	Extraction	0.93524	Face	image	outline
			562	204	182
2	Usage	0.93524	Face	image	extraction
			209	103	86
3	Result	0.93524	Face	extraction	outline
			158	105	63
4	Analysis	0.92325	Face	image	extraction
			64	30	29
5	Area	0.911563	Face	extraction	image
			275	155	143
6	Comparison	0.911563	Face	extraction	image
			82	49	41
7	Research	0.900169	Face	image	extraction
			89	50	48
8	Application	0.900169	Face	method	Extraction
			86	60	39
9	Inclusion	0.889056	face	image	Extraction
			82	47	36
10	Experiment	0.878214	face	image	Method
			64	47	27
11	Performance	0.878214	face	image	Method
			51	32	25
12	Matching	0.878214	image	face	Recognition
			27	24	24
13	Existing	0.867633	face	extraction	Characteristic
			63	34	29
14	Suggestion	0.857304	face	method	Image
			145	74	66
15	Transformation	0.857304	face	image	Extraction
			44	33	18

The words with the highest closeness centrality more than 0.85 closeness centrality value without the words in the first threshold value, derived from the network analysis, are designated as 'main keywords' in each individual period. For more accurate research trends analysis, co-occurrence matrix composed of main keywords in row and common keywords in column with a cell counting the occurrences between main and common keywords are needed to examine the relationship. Of all the common keywords, the three most frequently co-occurred common keywords—'relative keywords'—are extracted in proportion to their frequencies. Based on the relationship between main keywords and their relative keywords, this article is to analyze the research trends in biometrics.

IV. RESULTS

The result of the research trends in the fingerprint recognition is as follows: from 1990 to 2001, the extraction of characteristics of fingerprints, the matching between fingerprints, the recognition systems were main subjects. Additionally, the encryption of fingerprints' information was researched; from 2002 to 2008, the image and information of fingerprints were given much attention, and the characteristics of fingerprints were second main subjects. Other subjects included the preprocessing, the user authentication, and the matching technologies; from 2009 to 2015, the recognition system was a main subject. Other subjects included the matching technology, the user authentication, and the security of fingerprints' information.

The study of the research trends in the iris recognition shows the followings: from 2002 to 2008, the most attractive subject was the extraction of iris area. Other subjects included iris information, and the extraction of the characteristics of iris; from 2009 to 2015, much attention was paid to the extraction and application plan of iris. The other subject was the way of recognizing iris information without being affected by external environment like lighting.

The result of the research trends in the face recognition indicates followings: from 1990 to 2001, the extraction, usage, and real-time detection of the characteristics of a face were mainly focused as subjects; from 2002 to 2008, the extraction, usage, and real-time detection of the characteristics of a face were mainly focused as subjects as well. Other than these subjects, the management of the database about the characteristics of a face, the detection of a face area, the way of recognizing a face information without being affected by external environment like lighting, and PCA (principal component analysis) were popular as well; from 2009 to 2015, various subjects were focused on such as the detection of a face area, the way of coding a face area, analysis of the characteristics of a face, and the treatment of a face image.

The result of the research trends in the speaker—speech—recognition is as follows: from 1984 to 1989, the topics of the phoneme recognition, the voice activity detection, the recognition rate, the awareness of the speech section, the algorithm for speaker recognition, the comparison between voices prevailed; from 1990 to 2001, the most popular subjects were HMM(Hidden Markov Model), the phoneme recognition, the comparison and evaluation of the voice recognition performance, the independent speech recognition. Other subjects included the way of recognizing speech in a noisy environment, the continuous speech recognition, the word recognition; from 2002 to 2008, the main subject was the phoneme recognition, the recognition rate, the extraction of the characteristics of voice. Other subjects included the improvement of the existing voice recognition performance, the environment and noise about voice recognition, the extraction of the characteristics of voice, independent speech recognition, and HMM(Hidden Markov Model); from 2009 to 2015, the way of recognizing speech in a noisy environment, the database about voice, the

comparison between technologies, the improvement of technologies, the extraction and usage of the characteristics of voice were studied.

REFERENCES

[1] L. M. Sick, "A study on the fingerprint recognition system - focused on the trend of fingerprint recognition technology and the goal of policy," Master's Thesis, Dongguk University, Seoul, 2003.

[2] Y. M. Yoon, "A study on the quality evaluation method of biometric software — Focused on the evaluated cases of fingerprint recognition software," Doctoral Dissertation, Hoseo University, Asan, Korea, 2006.

[3] H. S. June, "A study on biometric standardization for biometric performance improvement," Master ś Thesis, Korea Universit, Seoul, 2001.

[4] S. M. Yun, "A study on automated immigration clearance system basedon biometric recognition technique," Master's thesis, Korea Aerospace University, Goyang, 2012.

[5] KISA, Survey for Information Security Industry in Korea: Year 2015. KISA, 2015.

[6] H. Lee, "Research trend analysis of Korean history subject area using CO-word analysis," Master's thesis, Yonsei University, Seoul, 2015.

[7] H. J. Moon, M. H. Lee, and K. H. Jeong, "Authentication performance optimization for smart-phon based multimodal biometrics," *Journal of Digital Convergence*, vol. 13, no. 6, pp. 151-156, 2015.

[8] B. Müller, "Security, risk and the biometric state: governing borders and bodies," London; New York: Routledge, 2010.

[9] A. K.J ain, K. Nandakumar, and A. Ross, "50 years of biometric research: Accomplishments, challenges, and opportunities," *Pattern Recognition Letters*, 2016.

[10] N. Clarke and S. Furnell, "Biometrics – The promise versus the practice," *Computer Fraud and Security*, vol. 9, pp. 12-16, 2005.

[11] S. Furnell and K. Evangelatos, "Public awareness and perceptions of biometrics," *Computer Fraud and Security*, vol. 1, pp. 8-13, 2017.

[12] J. Cho, "A study for research area of library and information science by network text analysi," *The Journal of Korea Society for Information Managament*, vol. 28, no. 4, pp. 65-83, 2011.

[13] H. J. Kim, N. K. Jo, and K. S. Shin, "Text mining-based emerging trend analysis for the aviation industry," *J Intell Inform Syst*, vol. 21, no. 1, pp. 65-82, 2014.

Jinhyo Jung is pursuing PhD in security convergence at Chung-Ang University, Seoul, South Korea. He obtained master's degree in security convergence at the same University. His research interests include Industrial Security, Physical Security, Information Security, and Security Convergence.

Hyunjung Yang is a graduate student with major in security convergence at Chung-Ang University, Seoul, South Korea. Her research interests include Industrial technology protection, criminal psychology, and data mining

Chang-Moo Lee is a professor of the Dept. of Industrial Security at Chung-Ang University in South Korea. He received his B.A. in political science from Yonsei University in South Korea, and both his M.Phil. And Ph.D. in criminal justice (security management) from the City University of New York in the U.S.A.