

Study on the Context Awareness that the Order Search System in a Distributed Computing Environment

Jeongah Han^{1*} and Laiwon Seo²

¹Department of Information and Communication, Chungnam National University, Republic of Korea; hja1128@empas.com

²Department of Game Engineering, Pai Chai University, Republic of Korea; seo@pcu.ac.kr

Abstract

In this thesis, order search system in established multi-media search technology is the system that shows the search results that users wish to words (order) using LBP technology system. Generally, the shopping mall or search system which is established to date, consist of this system. This system is not just to show the simple search results, but also show same results that user enter words no matter what it is Korean or English. In other words, search word; whether the search word is “ 사과 ” or “apple”, it draws on same results. Order search system is not based on potentially predominant PC, but it is miniaturization and uses the smart phone to fall in their small size, performance in memory. The reason why is realized to this system is to possible to offer powerful service, according to development in technology and miniaturization in some devices. Besides, in addition, enlarged some more images gathered on one screen had brought one screen needed even if images are stored on multiple servers. As this technology is possible in the society, it is necessary for users to make the system that shows the results which users want, if they don't input the contents. Realization order search system is informational to veridical circumstance in hypothetical space. According to system, it can be people-oriented autonomous service applied to intelligent methods such as recognition, characteristics in circumstance, learning and inferences.

Keywords: Component, Context Awareness, Distributed Computing, LBP, Searching System

1. Introduction

Humans are quite successful at conveying ideas to each other and reacting appropriately. This is due to many factors: the richness of the language they share, the common understanding of how the world works, and an implicit understanding of everyday situations. When humans talk with humans, they are able to use implicit situational information, or context, to increase the conversational bandwidth. Unfortunately, this ability to convey ideas does not transfer well to humans interacting with computers. In traditional interactive computing, users have an impoverished mechanism for providing input to comput-

ers. Consequently, computers are not currently enabled to take full advantage of the context of the human-computer dialogue. By improving the computer's access to context, we increase the richness of communication in human-computer interaction and make it possible to produce more useful computational services¹.

The reason why it is possible in society is the development in IT technology and scientific technique, also the amount of information is rapidly upsized. Moreover, the speed in the information development is predominantly fast, it brings huge effects on daily life. Therefore, it is necessary system that show the search results correctly in given circumstances. It is important to achieve correctly

*Author for correspondence

wished results rapidly in the search parts, but it is necessary for users to show intuitive and desirable contents reasonably. In this thesis, it shows that the system brings the order results, User's suitably desirable ask, using data stored in lots of computers. In addition, In order to use context effectively, we must understand both what context is and how it can be used. An understanding of context will enable application designers to choose what context to use in their applications. An understanding of how context can be used will help application designers determine what context-aware behaviors to support in their applications.

2. Distributed Computing

Ready access to large amounts of computing power has been a persistent goal of computer scientists for decades. Since the 1960s, visions of computing utilities as pervasive and as simple as the telephone have driven users and system designers. It was recognized in the 1970s that such power could be achieved inexpensively with collections of small devices rather than expensive single supercomputers. Interest in schemes for managing distributed processors became so popular that there was even once a minor controversy over the meaning of the word distributed. As this early work made it clear that distributed computing was feasible, researchers began to take notice



Figure 1. Distributed computing system.

that distributed computing would be difficult. When messages may be lost, corrupted, or delayed, robust algorithms must be used in order to build a coherent (if not controllable) system. Such lessons were not lost on the system designers of the early 1980s. Production systems such as Locus and Grapevine wrestled with the fundamental tension between consistency, availability and performance in distributed systems.

Distributed computing system is that application developer can freely utilize data stream processing logic in application area that is needed to consecutive processing about atypical data such as texts and images rather than consecutive processing based on data model. Also, it is focused on providing distributed computing infrastructure, in order to provide the base offered expandability about the amount of data. Intel and YACHOO etc carry out the study on distributed computing system. In this thesis, it is proved that consecutive processing and distributed nod based on distributed computing system about distribution of unit and load distribution, division of data stream, combination of parallel processing, method of delivering data stream in business, combination of stream consecutive processing technique based on computing technology handled with distributed nod obstacle is proved. Data stream processing calculation based on data model is used in application field in distributed computing system. IBM and DBMS business offer Info Sphere Streams distributed stream computing technique for expandability supply and flexible combination of application logic about increment of stream data. Info Sphere Streams is the distributed stream computing system offered data stream processing calculation based on distributed stream processing infrastructure².

3. Distributed Computing and Data Processing

Many analysts emphasize on careful decision to utilize Hadoop technique. In this case, it is necessary to extensive operation knowledge and analytics experts. However, most of global data center including IBM, GOOGLE and New York times don't have sufficient experts who are poor at using Hadoop technique connected big data. In these days, the amount of corporation using Hadoop technique increases, but it causes disappearance of vender support. Specifically, Hadoop technique is prior to video, audio and e-mail parts in comparison to traditional DBMS

(Data Base Management System). Also, it is advantageous to expand simply. The next is Hadoop technique solution instance³.

For example united states, 1% of all companies on December in 2012 use the Hadoop technique in production, manufacture and design fields, it increases to 2%~3% in the standard of last 2013. According to increasing data in geometrical progression, many companies choose the receptible. Solution carefully and need more enlarged combined data processing module, in order to meet consumer's requirements. In other words, more innovative customized data is necessary. Managing method is necessary to achieve needed decision making, because all task elements are appeared from consumer data's Value chain.

So as to support decision making progression efficiently, system to manage data and possible to control data is necessary from data occurrence to last consumer about information. To satisfy with quality of service delivery and decision making, to analyze original data correctly and partnership about data collection from various stakeholders is important. To enhance data management activity is necessary to achieve utmost decision making results that satisfy with requirements for all stakeholders. Also, users, process and investment management method of technology that maximize the data value is necessary to spread corporation's output in market circumstances efficiently^{4,5}.

4. Recognition Technique

Recognition technique means software and hardware technology that predict the user's ask in advance, as analyzing the user's life pattern, bio-signal and macro-environment and outputting qualified function situationally. Recognition technique is not only to appear all data, but also to interact with users and macro-environment. Also, this technique is cutting edge technology that added to concept of artificial intelligence that is predictable and analyzes the utmost decision making to satisfy with user's ask. For instance, it is possible to recommend purchased goods at the same time subdividing customer base aim to business marketing. In the case of using medical and public fields, it warns the cause of disease, disaster and crime. Recognition technique consists of the next three things largely.

First of all, network territory that sends the sensor and information which collect circumjacent context, secondly middleware that collects the sending situation informa-

tion and connects to proper information system, thirdly context engine that analyzes and infers to information in order to provide the answers in respective situation. If collected information by recognition sensors such as circumjacent GPS, RFID etc. uploaded information that send into network by context engine provide outputted result supply service based on "analysis, judgment, fusion, probability and inference rule about situation". Recognition service system consist of context interme-

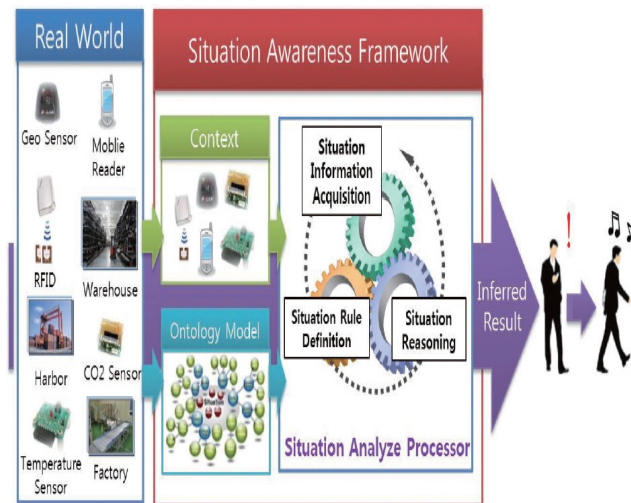


Figure 2. The knowledge engineering review, Cambridge Univ Press.

diary, device that recognize context, service agent that interact in situation and context information provided in intelligence situation⁶.

As universal use of smartphone in domestic area and ubiquitous environment spread, instances for using user navigation and recognition techniques are increasing. As used the navigation and location recognition application that can download from the smart phone recently, it shows the sudden development of location information and recognition that updates real-time location information and traffic situation and meet cloud technology to help the burden of hardware. Galaxy S4 or Galaxy Note³ phone see the technology based on S3 (health-care application) in Samsung. Developed Context-Aware Middleware for URC System (CAMUS) for Ubiquitous Robot Companion (URC) exists. CAMUS that helps u-Robot based on network can perform the recognition perform the recognition engine device that direct the work to managers, according to sensor that analyze the collected information, store and occur even⁷.

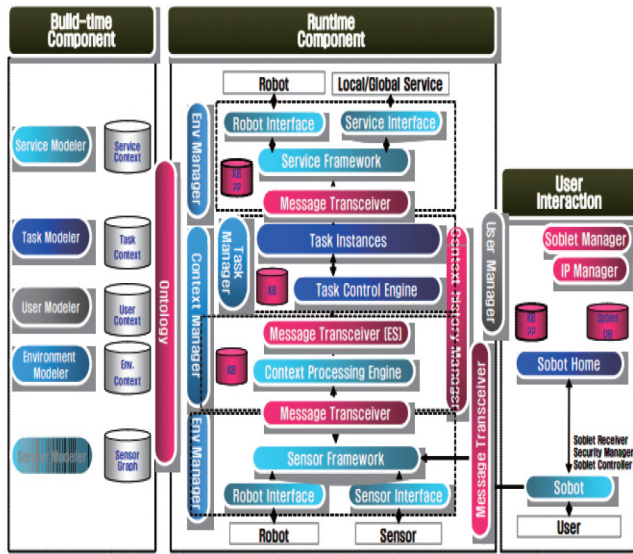


Figure 3. CAMUS: a middleware supporting context-aware services for network-based robots.

5. Recognition Technology and Search System

Meta data technology structure from images is divided into two parts such as Semantic DS that can express image's meanings, and Meta Information DS that can express typical data. Semantic DS writes 5W1H (When, Where, Why, What, Who, How) and Event in order to express specific information about object that images express, including specific information and condition. In the other way, Meta Information DS writes about simple explanation that is image's typical data technology structure, in comparison to Semantic DS, it is objective and includes image's copyright and producer's information and image⁸.

Figure 4 is XML Schema Diagram for Meta data. Respective images as mentioned are divided into seven

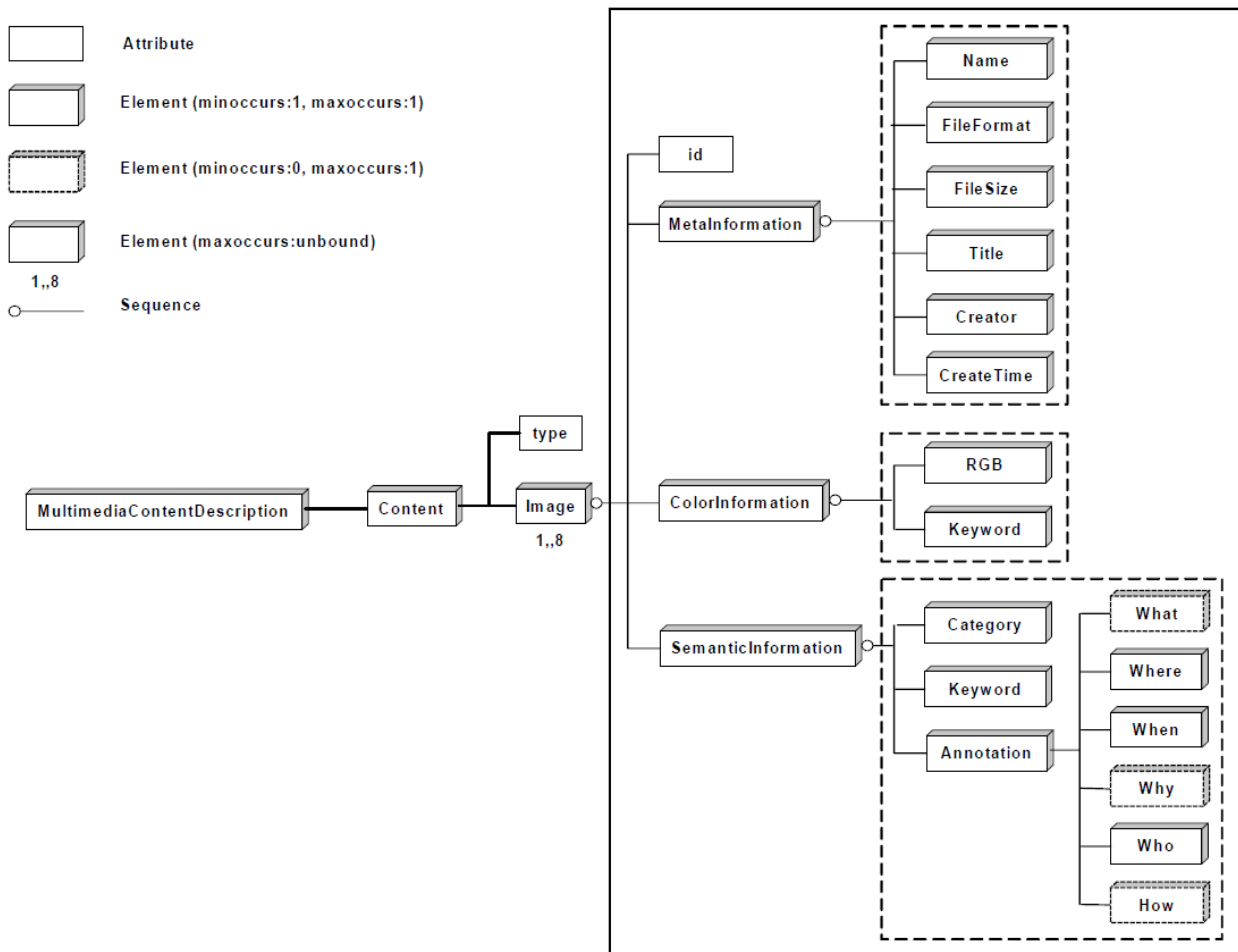


Figure 4. Image MetaData XML schema diagram.

parts such as Human, Animal, Plant, Art, Sport, and Architecture. Landscape in general division system. Respective image constitutes Physical information, Meta information, Color information and Semantic information.

Convergence framework is the way to provide combined service using one overlapped data model, unlikely method of using each other data model respectively application programs. Each application program, mobile terminating, access network and common model about much of information related to user is needed for fulfill of convergence framework. Also, development of protocol and architecture that is predictable exchange for each policy and information is followed by mobile terminating, user's past & present situation information, user's preference, and user⁹.

Like this, as analyzing the common collected data, Information Model for achieving information and knowledge and to make the search system based on this information is the this thesis's object. Situation recognition technique is generalized, like expert's anticipation, as sudden development of artificial intelligence algorism that finds concealed intention that analyzes and studies the huge amount of big data¹⁰.

The goal of context information acquisition should be to determine what a user is trying to accomplish. Because the user's objective is difficult to determine directly, context cues are used to help infer this information and inform an application on how best to support the user. Context awareness represents a generalized model of input (both implicit and explicit), allowing almost any application to be considered more or less context aware insofar as it reacts to input and the environment.

However, there is divergent opinion as to whether context should only comprise automatically acquired information or also include manually acquired information. In an ideal setting context would be obtained automatically and there would be no need for manual acquisition. However, in the real world not all context information can be sensed automatically and applications must rely on the user to provide it manually.

Context-aware applications are often distributed because they acquire context information from a number of different sources. 10 As much as the models for application distribution are well known, they are not always appropriate for distributed context information acquisition. Indeed, context awareness is most relevant when the

environment is highly dynamic, such as when the user is mobile.

Thus context-aware applications can be implemented on very diverse kinds of computing platforms, ranging from handheld devices to wearable computers to custom-built embedded systems¹¹. As a result context-aware applications require lightweight, portable and interoperable systems that can be implemented across a wide range of platforms.

Most present developed IT devices that utilize smart-phone and RFID devices interlocked IoT devices are developed. It proved that predictable and responding paradigm is generalized to enterprise, as industry field related to IT starts and spreads to all industries. Also, the development speed of changing suddenly mobile and correspondence environment will even be quicker than before. According to overall industries, try to do target marketing and analyze consumer's data using situation

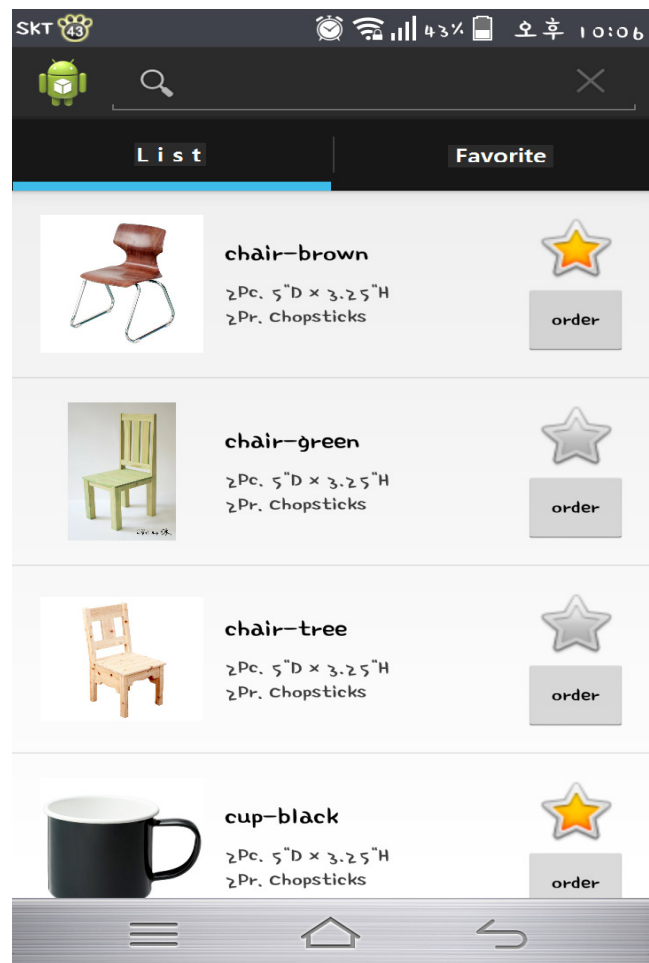


Figure 5. Order search system.

recognition service. Using the generalization of technology, to recognize and prove the hardware's devices, human beings voice, expression and body information is helpful for user to make the comfortable tool. Search system suggested by this thesis is not only to use the calculation skill, but also to develop inference function, and become the tool to interact each other in comfortable way at the same time figure out the user's intention.

6. Suggestion System

While source data is distributed, search function don't consider in the case of concentrating on data in the center. Distributed search situation is divided two parts. Data Fusion is to integrate results after distributed search for same data like Meta search in web, and Collection Fusion is to perform choice of search data source, change of quality and combination of search results, if there is other data in distributed storing place. Data Fusion is generally to normalize the range of suitability mark from respective source, because suitability mark exists in searched results from respective source. After search result from other

source and comparison is possible, the method of combining rank is used. Statistic and search method about each collection are different, because result combination in Collection Fusion is different to original of data. Thus, how to calculate the probability including each collection question & answer related proper document rather than how to normalize the mark range in result combination is usually used¹².

In this thesis, situation recognition search system materializes the order system that shows fast search results like search engine, according to search word that user inputs various goods image in many connected PCs. System is not to search the input words in the first time, but to bring the wanted results that search any words that user wants, regardless of storing location.

As the increment of information, the time of spending search is more increasing. In order to solve the problems and extract indexing of efficient information, categorizing about annotation information is performed. Index word n_i is the number of image that shows k_i in the total number of image in system. If the frequency of word k_i is $\text{freq}_{i,j}$ in the image, the frequency of k_i word is followed by (1)

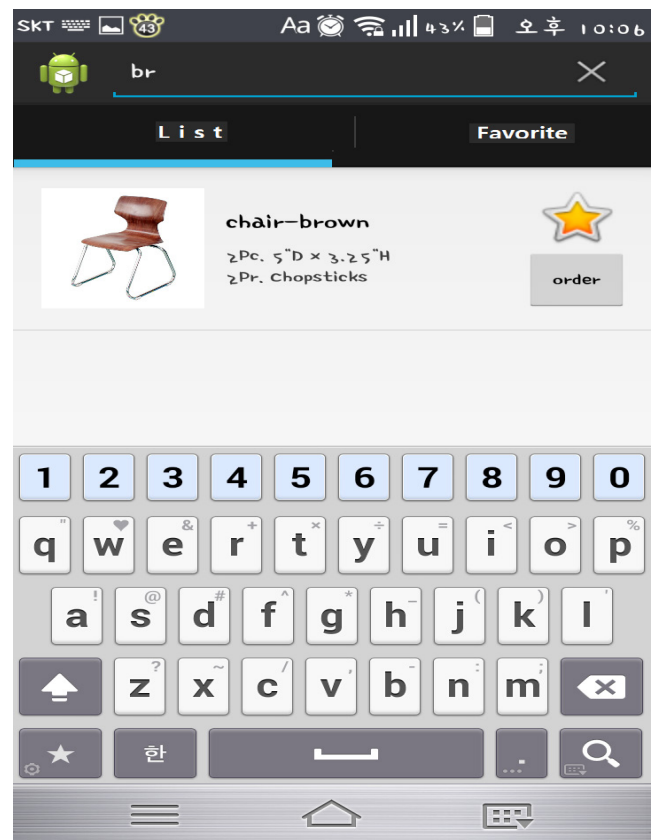
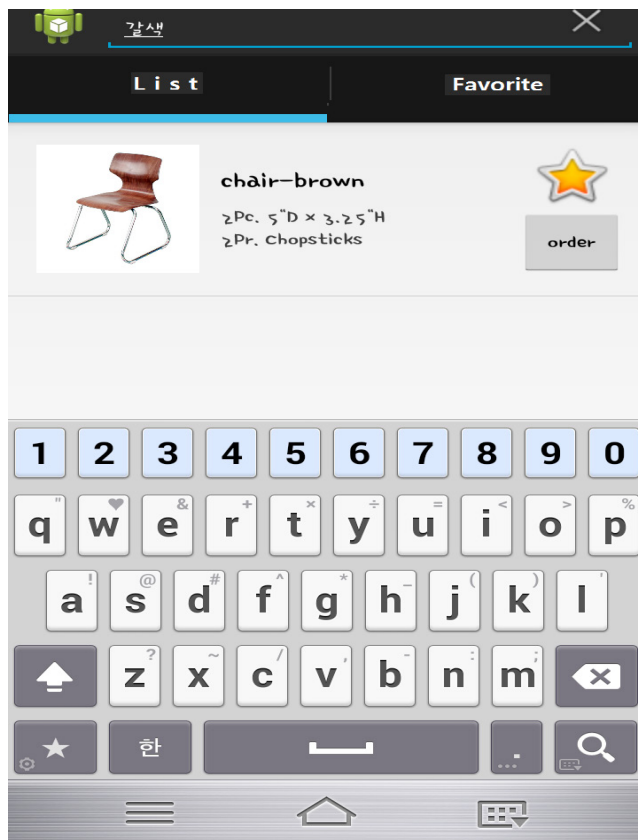


Figure 6. Found that the conditions according to circumstances.

in the image d_j^{13} .

$$f_{i,j} = \frac{freq_{i,j}}{\max_i freq_{i,i}} \quad (1)$$

The ceiling value in (1) appears to calculate all words in the text of image. If word k_i disappear in the text, it means $f_{i,j} = 0$. According to opposition of frequency about k_i , weight of word defines (2).

$$W_{i,j} = \alpha (f_{i,j} \times \log \frac{N}{n_i}) + (1 - \alpha) \times K_{i,j} \quad (2)$$

$K_{i,j}$ is keyword ceiling value in text and α is word's comparative importance. Ceiling value of word and importance of keyword ceiling value deducted in categorizing is same percentage as α to 0.5. Distributed network environmental situation recognition system suggested by thesis can be found system using LBP technology in existing goods image search techniques. It is as well as searching same artist's work efficiently, however searching original image into small image is not expected. Like this, experiment result will play predominant role in deciding the application range of search based on contents, and LBP system is impossible to content search in overall image range.

According to development of IT technology and scientific technique, the amount of information is increasing. In addition, the speed of information development is surprisingly fast and the effects of informationization in daily life are huge. In other words, more and more fast, correct and eidetic search engine is necessary. Many studies about image search method are progressing, because image search is important in internet search.

LBP is simple and wide in the application range. After experimenting system using this LBP technology, order image can be found in target image successfully. To search that LBP becomes image content search in overall range is difficult. Simply, some search words matching in light and simple system can be checked.

7. Conclusion

According to thesis, situation recognition search system in distributed computing environment can be checked image file search that user wished at the same time progressing rapidly. All multi-media file is not always fast,

because result is deducted by sampling file. However, according to situation recognition that user wanted, situation recognition search performance can be studied. According to study, searching file search using LBP in distributed computing environment can be checked by content search more rapidly.

This paper surveys seminal literature to better understand context awareness in mobile computing. It explores and discusses how context is represented, classified and applied and also discusses design principles with regard to context-awareness models. It also examines existing context-aware systems and frameworks supporting development of context-aware applications, and security and privacy issues with future developments in the research area.

Of course, all multi-media file is not always fast, because result is deducted by sampling file. However, searching multi-media file rapidly even in distributed computing environment is proved. According to development of this study, to improve multi-media search performance using smart phone, CCTV, RFID and sensor network will be studied.

8. References

1. Anind KD, Gregory DA. Towards a better understanding of context and context-awareness. 1999.
2. Rea R, Mamidipaka K. IBM info-sphere streams, enabling complex analytics with ultra-low latencies on data in motion. 2009.
3. Hadoop, the introduction of risk professionals are not ready to point out. CIO; 2012. p. 2.
4. Seok KY. The importance and utilization of Big Data-based predictive analytics. IDG Summary Performance 2012. CIO Korea; 2012. p. 12.
5. Big Data Social, Industry News. Industry Research; 2013. p. 2.
6. Gwon KH. Current and future context-aware computing. Vetaresearch & Consulting VETA Report. 2010; 10.
7. Moon SH. That the low status Computing Technology. Win-Win Tech; 2014. p. 6.
8. Rho S, Hwang E. Schema design and implementation for multimedia data retrieval. 2002; p. 10.
9. Autonomic Communication. Available from: <http://www.autonomiccommunication.org/>
10. Hyun KK. Situational awareness technology in medical diagnostic industry. Fusion And New Growth Industry Webzine. 2013; 16:6.

11. Bauer M, Heiber T, Kortuem G, Segall Z. A collaborative wearable system with remote sensing. ISWC'98. Proceedings of the 2nd International Symposium on Wearable Computers; 1998 Oct 19-20; Pittsburgh, PA. Washington: IEEE Computer Society; 2000. p. 10–7.
12. Ricardo BY, Berthier RN. Modern Information Retrieval. ACM press; 1999.
13. Hyeon MS. Current status and future direction of information retrieval technology. Journal of Computing Science and Engineering. 2004; 4.