Trend Surfing System based on Multi-aspect, Multi-screen and Multimodal User Interface

Toru Kobayashi
Division of Electrical Engineering and Computer Science
Graduate School of Engineering, Nagasaki University
Nagasaki, Japan
toru@cis.nagasaki-u.ac.jp

***Abstract*—The multi-screen system that contains the advanced interactive user interface has attracted a great deal of attention. The advanced interactive user interface contributes to the realization of a more personalized information usage environment. However, this interface causes the problem of unfair information retrieval to users who do not have the ability to make sufficient use of information technology. Therefore, this paper proposes a "Trend Surfing System" that enables the closing of the information retrieval capability gap. In this proposed system, users can automatically pull up relevant information (multi-aspect) one after another just by pointing the desired trendy information listed on a screen device. Users can also simultaneously distribute relevant information to different screen devices (multi-screen) around users by tapping, motion or voice operation (multimodal user interface). In order to enable the multi-aspect information retrieval, we propose a two-phase information retrieval model. We also propose the system architecture which enables double mashup not only web services but also screen devices in front of users for the multi-aspect information retrieval and multi-screen. In terms of multimodal user interface, we apply for Web standardized technologies. We show the two-phase information retrieval model for accomplishing the proposed system, system architecture and user interface. We also describe the outline of a prototype system implemented as Web application and evaluation results. Finally, we discuss the efficiency at the point of closing the information retrieval capability gap based on the prototype evaluation results.**

Keywords— information retrieval; multi-screen; HTML5; Service-Oriented Architecture; user interface

I. INTRODUCTION

Thanks to the spread of useful, high-performance screen devices like tablet computers or smart phones, multi-screen systems containing advanced interactive user interfaces have gained attention. In particular, research and development is aggressively progressing in the area of IP TV[1][2][3] or Smart TV[4] which has merged broadcasting and communication. Although these multi-screen systems relating to TV media have different features, they have one feature in common: an advanced interactive user interface. Through this interface, users can perform more active information retrieval synchronized with user's preferences. The advanced interactive user interface contributes to the realization of a more personal information usage environment.

On the other hand, this active information retrieval environment causes the problem of unfair information retrieval to users who do not have the ability to make sufficient use of information technology. For example, let us consider two user groups. The first user group consists of users who are accustomed to actively using tablet computers or smart phones to get information every day without problems. The second user group consists of users who normally get information only passively, through broadcasted TV. If both groups used the multi-screen system equipped with the advanced interactive user interface, the retrieved information results might not be the same. It is obvious that the first user group would retrieve richer information than the second user group. In brief, the information retrieval capability difference causes the gap[5][6].

Therefore, this paper focuses on the issue of the information retrieval capability and proposes a "Trend Surfing System" that enables the closing of the information retrieval capability gap. Trend surfing in this paper means that some kinds of related trendy information can be displayed on different screen devices according to the user's direct operation. We call these some kinds of related trendy information "multi-aspect" information. The appearance of such an action would be like net surfing across different screen devices. The target information of this paper is current trends, such as hot words, news topics, and SNS topics. Hot words mean popular searching keywords on search engines like Google. Hot words reflect the interests of a majority of society. News topics are the information distributed from the mass media. SNS topics indicate posted messages among friends of the target user. This trendy information might be new to users at any time, so the information retrieval capability gap could be observed more easily. For this reason, this paper focuses on trendy information.

The proposed system consists of several screen devices such as tablet computers. For example, hot words, news and SNS topics will be listed just after launching a browser on a tablet computer. After selecting one of the listed items, not only the detailed information but also the different aspects of related information, such as thumbnails of social movies or private movies, tweets, book or product summaries, will be shown. Related movies can be displayed on different screen devices by a drag and drop action on a tablet computer, by shaking a tablet computer or by speaking words. Detailed information about news topics, SNS topics, and products also can be shown on a tablet computer. Figure 1 shows this usage image. As indicated in Figure 1, we call such environment consisting of some devices equipped screen functions the "multi-screen" environment. In this environment, one screen device should be assigned as an operation device by users like a tablet computer in Figure 1. We call it a conductor screen device. The other screen devices will be used for showing multi-aspect information. We call them performer screen devices.

From the above, the proposed system makes it possible for users to discover trendy information without inputting any keywords. This system automatically collects the different aspects of related information instead of users. This system also enables users to distribute the related information on a multi-screen device environment according to the user's intuitive operation that means a multimodal user interface. This system will help users to understand the meaning of the retrieved trendy information. We can say that this system will have the possibility of the realization of a fair information retrieval environment which can be used for closing the information retrieval capability gap.

In order to enable the multi-aspect information retrieval, we propose a two-phase information retrieval model which contains the first contact and the second contact from users. We also propose the system architecture which makes it possible the double mashup not only web services but also screen devices in front of users for the multi-aspect information retrieval and multi-screen. In terms of multimodal user interface, we applied for advanced Web standardized technologies. We also implemented a prototype system as Web applications for evaluation.

We show the related work in Section II. We define the requirements and the two-phase information retrieval model in Section III and explain the implementation method for the Trend Surfing System in Section IV. Then, we show the prototype system outline and evaluation results in Section V. Finally, we discuss the advantage of the proposed system in Section VI.

II. RELATED WORK

The purpose of this paper involves "multi-screen" and "information retrieval," so we describe those two issues as related work.

A. Multi-screen

Multi-screen systems in the area of TV can be classified into three cases [2]. The first case is when users watch contents using screen devices having different specifications. Different specification examples are the screen size, CPU power, codec, and network bandwidth. The second case is when users seamlessly watch contents over different devices. For example, a user watches contents using a smart phone while commuting. After getting home, this user resumes watching the same contents using a wide screen TV from the interrupted point. The third case is when users watch main contents, like movies, on a main screen device and also simultaneously watch other related contents on different screen devices. For example, when a user is watching a movie scene on a main screen device, advertisements of clothes equivalent to one of an actress are shown on the other screen devices[7].

Our study approach more closely involves the third case. Normally, the third case includes story content, such as movies-oriented content, because other related content will be

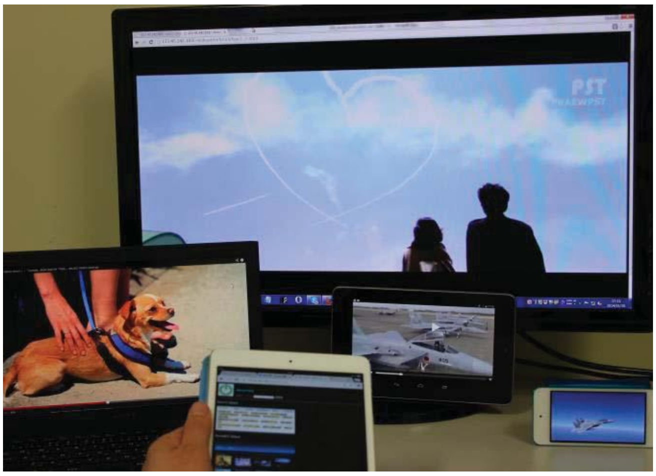

Figure 1. Multi-screen Usage Image

modified by the story content's progress. In this case, users can not take part in the story content's progress. On the other hand, in our study, the main screen device's content is trendy information items that can be dynamically selected by users. Then, the other aspects of related content will be displayed on the other screen devices. In this way, we can say that our study approach is the most oriented towards trendy information. In existing studies categorized into the third case, only information related to the story content can be displayed on the other screen devices. However, in our study, users can actively select the desired, trendy information items to see more closely related information, which will be displayed on the other screen devices. That is the difference between existing contributions and our study.

B. Information Retrieval

Currently, there are some information retrieval studies that involve giving users several kinds of related information without relying on the user's information retrieval capability. For example, Memorium[8], Discover Apps[9], Goromi-web[10], and InfoSkin[11] are trying to give users some kinds of related information in a passive way.

In Memorium, cards describing keywords move autonomously inside a screen. When one card bumps into another card, the new card describing new keywords will be created. Memorium provides a client application that can be used for browsing detailed information. However, Memorium requires initial keywords. Creation of a new card describing new keywords depends on a random system operation, so that users can not reflect their intentions in the process of new card creation.

Discover Apps can be used for searching iOS applications in a social graph manner. Candidate application icons will emerge after keywords are given. When an icon is tapped, detailed application information appears. Users can also reach similar applications by tracing a social graph. Goromi-web achieves passive web browsing. Categorized web pages appear one at a time, like clouds drifting in the sky. Users can be involved in deciding the drifting direction. This system makes it possible for users to discover fortuitous information through minimum user interaction. These studies enable users to passively retrieve related information by reflecting the user's intention. However, they need initial keywords.

Infoskin focuses on a window shopping metaphor in the real world. Relevant information items indicated as icons move in a screen. Users can tap or drag and drop these icons to see detailed information or reach different aspects of related information. Infoskin does not compel users to input search keywords because it allows the use of hot words as initial search keywords. However, although Infoskin is a user interface that can be used for passively retrieving related information, it has not been adapted for the multi-screen environment nor the multimodal user interface.

The features of the proposed system are the keywords involving less searching at the initial stage, reflecting user's intentions against the information retrieval direction, multi-screen usage, and multimodal user interface availability. Realizing these four points simultaneously is completely different from the existing work.

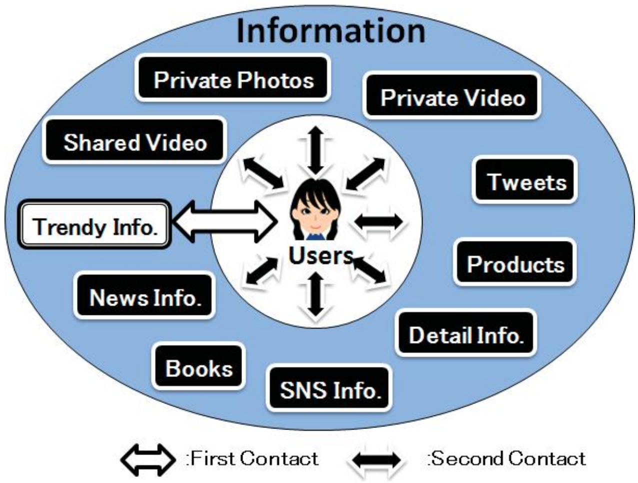

Figure 2. Two-phase Information Retrieval Model

III. REQUIREMENTS AND INFORMATION RETRIEVAL MODEL

In this section, the requirements for the proposed Trend Surfing System are described. Then, the information retrieval model for satisfying the requirements is described.

A. Requirements

The research object involves the three requirements listed below.

- Requirement 1: to avoid the "cold start problem," which concerns the issue of when the system cannot draw any inferences for users about which it has not yet gathered sufficient user information.

  For example, users can see trendy information without inputting search keywords.

- Requirement 2: to realize the different aspect information retrieval by reflecting the user's intentions.

  For example, users can select a desired item from the presented trendy information items. Then, different aspects related to the information of the item selected by users are presented. Users can also choose one of such related information items.

- Requirement 3: to show the different aspect information based on a multi-screen environment via multimodal user interface.

  For example, the different aspects related to the information can be placed on the multi-screen environment according to the user's intuitive operation.

B. Two-phase Information Retrieval Model

Figure 2 shows the proposed two-phase information retrieval model. At first, users, indicated in the middle of Figure 2, can see the trendy information items such as hot words, news topics, and SNS topics. In this paper, we call this action the first contact. Users can select a desired trendy item, then different aspect related information candidates will be presented. Users can also choose one of the candidates to retrieve more precise information. We call this action the second contact.

In this model, users do not have to input search keywords to find trendy information at that time. Users only have to view and select one of the presented information items. In this way, users can retrieve the details of the desired trendy information from several kinds of aspects. Currently, users need to think about search keywords and individually input them into web services. However, in this model, users do not have to think of search keywords nor input them separately into web services.

IV. IMPLEMENTATION

In this section, we explain how to implement the Trend Surfing System based on the two-phase information retrieval model.

A. System Architecture

Figure 3 shows the system architecture of the proposed system. The main feature of this system architecture is that the central control server performs two kinds of combinations (mashup) due to Service Oriented Architecture. For example, the control server establishes the combination for external web services using the open application interface (API). The control server also combines several screen devices such as a conductor device and performer devices in front of users. The protocol between the control server and each screen device is HTTP and WebSocket[12], so that heterogeneous screen device combinations can be available just by supposing the browser installing into each device. In this way, the control server allows users to enjoy a multi-screen environment consisting of heterogeneous screen devices surrounding users.

B. System Configuration

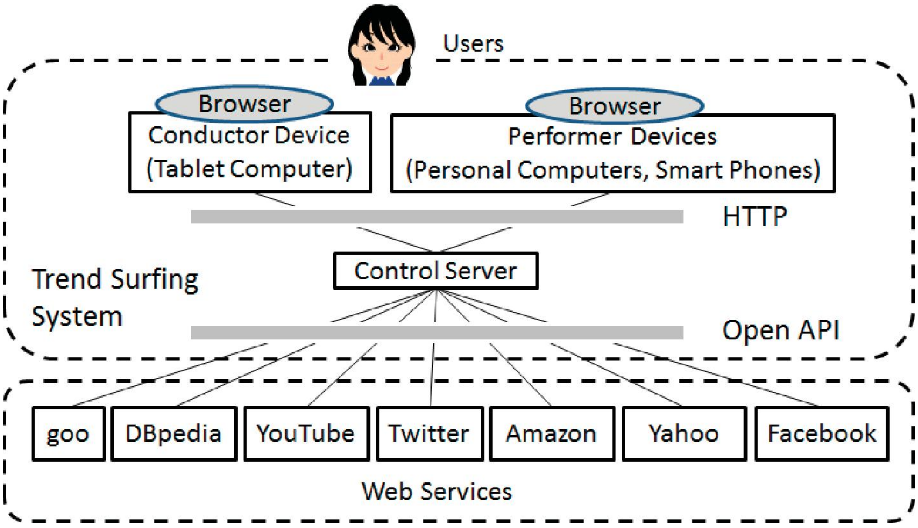

Figure 3. System Architecture for double mashup

Figure 4 shows the system configuration of the proposed system. Screen devices operated by users are classified into a conductor screen device (Conductor S.D.) and performer screen devices (Performer S.D.). The conductor screen device is mainly operated by users. The performer screen devices are used to show many kinds of information. Therefore, on the control server, the server application as well as the HTML resources for the conductor screen device and the performer screen devices (Operation AP and Display AP, respectively) has been arranged.

The server application uses WebSocket to communicate with the conductor screen device or the performer screen devices. The server application also has the Get Web Information function to obtain information from external web services. The conductor screen device and the performer screen devices allow HTTP access to the control server to obtain HTML resources, such as the operation application (Operation AP) and the display application (Display AP), respectively, via each browser. Each HTML resource includes WebSocket implemented JavaScript so that these screen devices and the control server can communicate with each other.

The operation application has the ability to call the Get Web Information implemented in the control server through HTTP. Therefore, the conductor screen device can follow two paths: information gathering via WebSocket as well as through HTTP. This enables a more flexible system configuration for gathering information. The operation application also has a multimodal user interface (Multimodal UI). This allows us to operate the conductor screen device using the drag and drop operation by a touch sensor, a motion operation by an acceleration sensor or a voice operation by speech recognition. The motion operation has been implemented using Device Orientation Event Specification [13] which is the acceleration sensor API for HTML5. The voice operation has also been implemented using Web Speech API Specification [14] which is the speech recognition API for HTML5. All these multimodal user interface functions have been implemented by JavaScript.

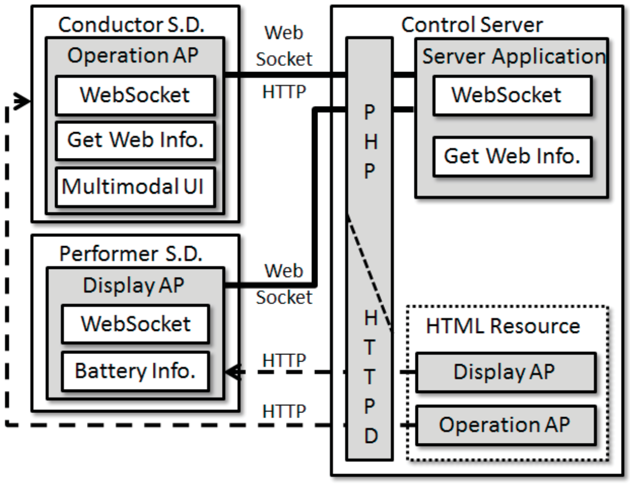

Figure 4. System Configuration

The display application has the ability to detect battery information. The performer screen devices might be a battery-operated device in some cases. Then, we implemented the function to regularly notify the operation application about the remaining battery quantity via WebSocket based on Battery Status API [15] by JavaScript.

C. Process Flow

The process flow of the proposed system contains three procedures, described below. The state transition diagram, which describes the process flow, is shown in Figure 5.

1) Gathering and displaying first contact items (1st Contact Items): After executing the operation application on the conductor screen device, the trendy information of the first contact items will be displayed. Specifically, the control server collects hot words from "goo", a Japanese search engine, news topics from Yahoo, and SNS topics from Facebook. goo, Yahoo and Facebook are providing the keyword ranking information as RSS [16], News Web API [17] and Facebook API [18], respectively. Concerning SNS topics, only the top 10 posted messages sorted in reverse order of "likes" were considered, because the number of "likes" might reflect the interest rate among friends of the target user. Then, these information items will be displayed on the conductor screen device as the first contact items.

2) *Gathering and displaying second contact items (2nd Contact Items):* When one of the first contact items is selected by users, a search keyword will be automatically generated to collect second contact items from the external web services. Then, the second contact items will be displayed on the conductor screen device. Specifically, the control server extracts a search keyword from the selected first contact information using the key phrase extraction Web API provided by Yahoo! Developer Network[19]. Due to this auto-generated search keyword, the control server will collect the second contact items from DBpedia, YouTube, Twitter, and Amazon using each open API[20][21][22][23] and display them on the conductor screen device.

3) *Gathering and displaying the selected, second contact item's detailed information (2nd Contact Information):* This process collects and displays detailed information about the selected, second contact items. Specifically, the control server distributes the second contact item's detailed information to the same tab or a new tab on the conductor screen device browser, or different performer screen devices' browsers. The output destination decision of the detailed information depends on its media type, like texts, pictures, and movies, because of the viewpoint of the user's visibility (Table 1). The different second contact items can also be redisplayed by selecting different first contact items.

D. User Interface

Figure 6 shows a design diagram of the conductor screen device user interface. The proposed user interface is divided into an upper part, called the first contact zone, and the bottom part, called the second contact zone. In the first contact zone, the first contact items collected by the control server are displayed according to the first procedure described in the previous section. On the other hand, in the second contact zone, the second contact items collected by the control server are shown according to the second procedure described in the previous section.

Almost all operations can be performed by voice as well as tapping or dragging and dropping using fingers or a mouse. If the conductor screen device contains an acceleration sensor, YouTube movies following the third procedure described in the previous section can be shown on the performer screen devices by shaking the conductor screen device. If there were some performer screen devices, performer screen device selecting turn would be determined according to the order of connecting time from a performer screen device to a control server via WebSocket.

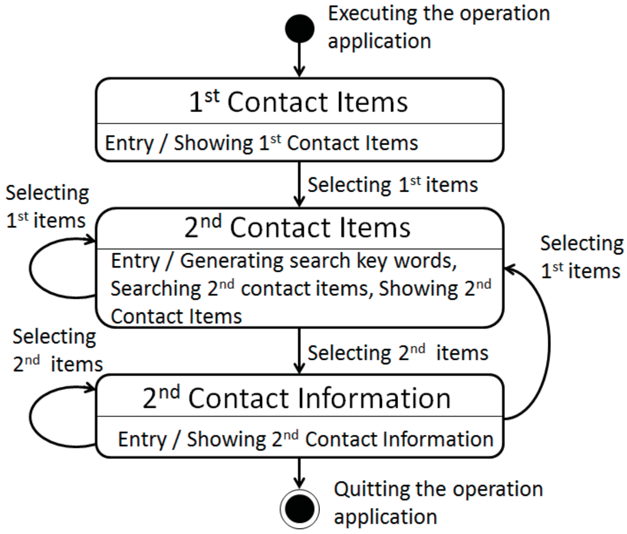

Figure 5. State Transition Diagram

Table 1. Output Destination Decision

Destination / Media Type	Conductor S.D.		Performer S.D.
	Same Tab	New Tab	
Text (DBpedia, Twitter)	○	—	—
Text, Pictures (Amazon, Yahoo, Facebook)	—	○	—
Movies (YouTube)	—	—	○

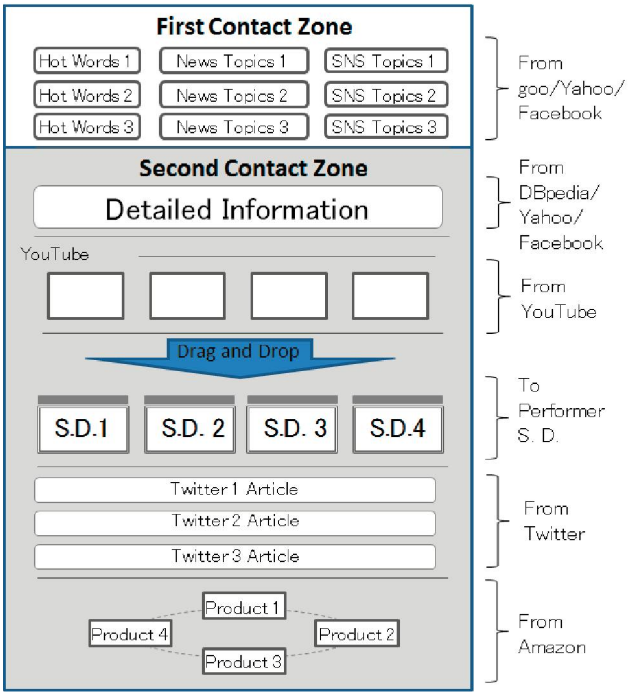

Figure 6. Design Diagram of Conductor Screen Device User Interface

V. Prototype System

We implemented a prototype system consisting of the control server, the conductor screen device, and the performer screen devices specified in Table 2. In terms of the control server, it was constructed in two ways. One was to set up the control server on the local area network of our laboratories. The other was to set up the control server on the Amazon Web Services (AWS)[24] cloud system. We confirmed that both ways were applicable for prototype system evaluation. Figure 7 shows the user interface image of the conductor screen device, and Figure 1 shows the multi-screen usage image.

We confirmed that the prototype system could extract the first contact items without inputting keywords. We also confirmed that the second contact items were changed within seconds to new ones just after tapping or speaking the first contact items on the conductor screen device as indicated in Figure 7. In this way, we could realize a change and educe different aspects related to the information in succession according to the user's successive tapping or speaking of the desired first contact items.

Figure 1 presents an example case in which the hot word relating to a Japanese popular TV drama was tapped as the first contact item. The different scenes could be shown on a Windows PC, iPod Touch, Nexus 7 as the second contact information. The promotional video of the TV drama theme song could also be shown simultaneously on a different performer screen devices (note PC in Figure 1). These operations could be completed by dragging and dropping or speaking the second contact items onto an iPad Mini or a Windows Tablet as the conductor screen device. Due to the performance of this kind of trend surfing by users, it has been confirmed that users could grasp different aspects of the TV drama's associated information, such as stories, heroes or heroines, and theme songs, even if users have never seen such a TV drama. In these results, we could say that all defined requirements in Section A of III have been satisfied.

Figure 8 shows an example of tapping a Facebook posted message as the first contact item. In this example, the auto-generated keyword, "three-dimensional switchback," has been extracted from the tapped Facebook posted message. A few YouTube movies searched by the keywords "three-dimensional switchback" could be simultaneously shown just by shaking the conductor screen device. In this manner, concerning user performance of this kind of trend surfing, not only could actual trains be displayed during live filming, but also toy train movies could be displayed simultaneously. By this experiment, it has been confirmed that users could correctly grasp the meaning of "three-dimensional switchback" in a short time.

Table 2. Prototype System Specification

Control Server	Conductor S.D. / Performer S.D.
・Linux CentOS 6.4 Apache 2.2.15 PHP 5.3.3 WebSocketServer 2.1.0	・iPad Mini, iOS 7.0.2, Safari ・Windows PC, Win7, 8, Chrome ・iPod Touch, iOS 7.0.3, Safari ・Windows Tablet, Win8, Chrome ・Nexus7, Android 4.3, Chrome

Figure 7. Conductor Screen Device User Interface Image

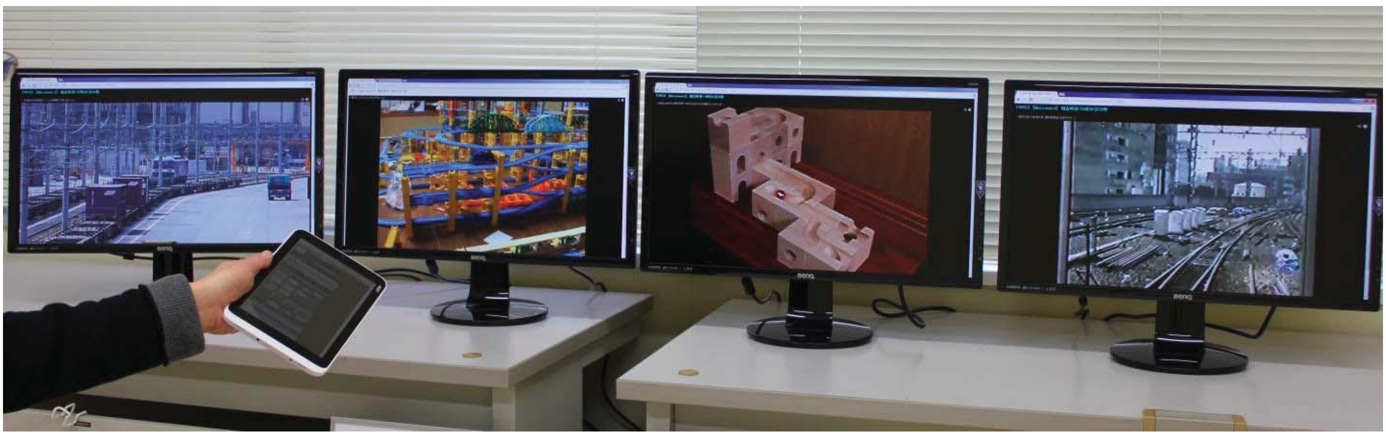
Figure 8. Trend Surfing Example

VI. DISCUSSION

This study focuses on the issue of information retrieval capabilities and aims to create a fair information retrieval system independent of user's information technology literacy. As the first step, we implemented a "Trend Surfing System" that targets trendy information. Until recently, the most popular method to retrieve trendy information was the newspaper. The general newspaper approach is listed in the steps below.

- Identifying desired news articles by looking over all headlines.
- Reading identified news article

These steps correspond to the first and second contacts described in the two-phase information retrieval model in this paper. In short, this proposed system follows the same pattern as information retrieval from newspapers. The newspaper is read all over the world and the problem of information retrieval capability differentials among users has not occurred. It is suggested that the proposed system can solve the problem of information retrieval capability differentials.

Moreover, in the case of newspapers, one headline corresponds to one article. However, the proposed system can simultaneously show different aspects of related contents on different performer screen devices via multimodal user interface. The cases of the popular Japanese TV drama and unknown keyword "three-dimensional switchback," mentioned in the previous section, present good examples. This means that the proposed system not only follows but also extends the newspaper's pattern of use. For this reason, the proposed system is expected to not only retrieve trendy information independent of user's information technology literacy but also understand the meaning of the information, itself.

On the other hand, Web technologies and Web services will be improved day by day. In order not to become the system obsolete, we need to apply excellent external technologies or services continuously instead of implementing them by ourselves. In this paper, we showed the system architecture based on Service Oriented Architecture. We also indicated the applicable system configuration and the process flow which applied Web standard specifications and also connected the external Web services. In this mean, the proposed system would be improved easily just modifying the part of new Web standard specifications or new Web service API if Web standard specifications or external Web services were improved or emerged as new ones. This could contribute the sustainability to the proposed system.

VII. CONCLUSION

This paper focused on the issue of information retrieval capabilities and then proposed a "Trend Surfing System" in order to close the information retrieval capability gap. In the proposed system, users can educe different aspects related to information one after another just after pointing out the desired trendy information item. Users can also distribute the related information to the different screen devices around users through an intuitive user interface like surfing across different devices.

The proposed system follows the same pattern as the information retrieval from newspapers read all over the world. Moreover, the proposed system can provide multi-aspect information through the multi-screen device environment operated by multimodal user interface. This means that this proposed system could be expected to not only retrieve trendy information independent of a user's information technology capabilities but also understand the meaning of the retrieved information, itself. The cases of the popular Japanese TV drama and unknown keyword "three-dimensional switchback" as results of prototype system evaluation can support this conclusion.

In this paper, trendy information has been targeted. In order to expand the target information, we will consider an advanced personal functionality, such as a context-aware technology. Then, we will try to merge it with the proposed system by making good use of flexible system architecture.

REFERENCES

[1] Peng Tan, Slevinsky, J., "Multi-screen IPTV: Enabling technologies and challenges,"Proc. 2011 IEEE International Conference on Consumer Electronics (ICCE), pp. 1-2, Las Vegas, USA, Jan.2011.

[2] Changwoo Yoon, Taiwon Um, Hyunwoo Lee," Classification of N-Screen Services and its standardization,"Proc. 2012 14th International Conference on Advanced Communication Technology (ICACT), pp. 597-602, PyeongChang, Korea, Feb.2012.

[3] Jiwon Jang, Hyunwoo Nam, Younghan Kim, "Mobile device-controlled live streaming traffic transfer for multi-screen services,"Proc. 2012 International Conference on Information Networking (ICOIN), pp. 415-420, Bali, Indonesia, Feb.2012.

[4] Ohmata, H., Takechi, M., Mitsuya, S., Otsuki, K., Baba, A., Matsumura, K., Majima, K., Sunasaki, S., "Hybridcast: A new media experience by integration of broadcasting and broadband," 2013 Proc. of ITU Kaleidoscope: Building Sustainable Communities (K-2013), pp. 1-8, Kyoto, Japan, April 2013.

[5] Kobayashi, T., "A Proposal of Smart TV System focused on Findability," 2013 IEEE 2nd Global Conference on Consumer Electronics (GCCE), pp.507-508, 2013.

[6] Kobayashi, T., "Information Orchestration System for Closing the Information Retrieval Capability Gap," Proc.2014 IEEE 38th Annual International Computers, Software and Applications Conference(COMPSAC2014), pp.644-645, 2014.

[7] Tae-Beom Lim, KyungWon Kim, Seok Pil Lee, Kyoungro Yoon, "Real-time adaptive advertising1 framework based on MPEG-21 for multi-screen IPTV environment, " 2012 IEEE International Conference on Consumer Electronics (ICCE), pp.13-16, Jan. 2012.

[8] Watanabe, K., Yasumura, M.," Memorium: The concept of persistent interface and its prototype," Workshop on Interactive Systems and Software, WISS2002,pp.99-104, Nov. 2002.

[9] Discover Apps, http://discovr.info/, available at 2013.05.31.

[10] Otsubo G.,, "Goromi-Web browsing for unexpected information on the web,"Proc. of the 6th ACM SIGCHI conference on Creativity & Congnition, pp.267-268, 2007.

[11] Seko S., Aoki R., Ihara M., Kobayashi T., "InfoSkin:User Interface for Passive Information Browsing," Proc. of 2012 Joint Conference of the International Industrial Information Systems Conference & the International Conference on Computers, Communications and Systems, pp.3-4, Chiang Mai, Thailand. Dec.2012.

[12] The WebSocket API, http://www.w3.org/TR/2011/WD-websockets-20110929/, available at 2014.05.08

[13] DeviceOrientation Event Specification, http://dev.w3.org/geo/api/spec-source-orientation.html, available at 2014.05.08

[14] Web Speech API Specification, https://dvcs.w3.org/hg/speech-api/raw-file/tip/speechapi.html, available at 2014.05.08

[15] Battery Status API, http://www.w3.org/TR/battery-status/, available at 2014.05.08

[16] Keyword ranking information as RSS provided by goo, http://www.goo.ne.jp/rss/, available at 2014.08.27

[17] News Web API provided by Yahoo! JAPAN, http://developer.yahoo.co.jp/webapi/news/, available at 2014.08.27

[18] Facebook API, https://developers.facebook.com/docs, available at 2014.08.27

[19] Key phrase extraction Web API, http://developer.yahoo.co.jp/webapi/jlp/keyphrase/v1/extract.html available at 2014.08.27

[20] DBpedia API, http://dbpedia.org/About, available at 2014.08.27

[21] Youtube API, https://developers.google.com/youtube/getting_started?hl=en, available at 2014.08.27

[22] Twitter API, https://dev.twitter.com/, available at 2014.08.27

[23] Amazon API, http://docs.aws.amazon.com/AWSECommerceService/latest/DG/Welcome.html, available at 2014.08.27

[24] Amazon web services, http://aws.amazon.com/, available at 2014.08.27