Multimedia Presentation Authoring and Virtual Collaboration in Medicine

Chuleui Hong, *Member, KIMICS*

***Abstract*—Web-based virtual collaboration is increasingly gaining popularity in almost every area in our society due to the fact that it can bridge the gap imposed by time and geographical constraints. However, in medical field, such collaboration has been less popular than other fields. Some of the reasons were timeliness, security, and preciseness of the information they are dealing with. In this paper, we are proposing a web-based distributed medical collaboration system called Virtual Collaboration System for Medicine (VCSM) for medical doctors that meet the needs.**

The proposed system consists of two parts – multimedia presentation and recordable virtual collaboration. The former supports synchronized multimedia presentation using Synchronous Multimedia Integration Language (SMIL.) It allows synchronization of the contents of a PowerPoint presentation file and a video file. The presentation may be provided to the participants before the discussion begins.

Next, in the virtual collaboration stage, participants can use text along with associated symbols during the discussion over the presented medical images. The symbols such as arrows or polygons can be set or removed dynamically to represent areas of interest in digital images using so called layered architecture that separates image layer from annotation layer. XML files are used to record participants' opinions along with the symbols over some particular images

***Index Terms*—Virtual Collaboration, SMIL, Real-time Digital Image Annotation, Multimedia Presentation Authoring.**

I. INTRODUCTION

In modern society, Internet-based communication has become an important commodity in daily life, anywhere from instant messenger to hi-speed video conferencing. The communication mechanism can be in many different ways such as text chat, messaging, audio and video. Internet-based collaboration or virtual collaboration may be carried out in various ways e.g., simple text exchange through web pages, text chat, voice-over-IP and IP video conferencing [1]. Olson et al. [2] stressed and analyzed collaboration among researchers and assessed some of their collaboration projects. The demand for the virtual collaboration has been overwhelming in the sectors including research community, academia, medical and some industry for various purposes, e.g., researchers getting better results by doing collaboration with their peers, schools offer distance education to serve broader community, medical doctors collaborate for better treatment of patients – by getting experts' opinions, industry use collaboration for more efficient sales and marketing meetings, etc. Suebnukarn et al. [3] proposed a tutoring system for medicine that makes use of chat, discussion and experts' opinion to draw a conclusion. Pourdadash et al. [4] proposed web-based medical image analysis tool using some image processing and web-based communication. Some systems regarding medical annotations were proposed. Pooshfam et al. [5] proposed a system for annotating images and videos in a collaborative way. In medicine field, such collaboration methods didn't get much attention due to some medical specific requirements that may not be satisfactorily met by currently available collaboration technologies. Some of the requirements can be:

- Accuracy of patient data
- Security: only the authorized person should have access to the data
- Time-stringent discussion
- Annotation on medical images

In this paper, we are proposing a new way of collaboration and education in medicine that satisfies the above requirements as well as a tool called virtual collaboration system for medicine (VCSM) that helps researchers or medical doctors exchange opinions over digital images with aforementioned requirements in mind. VCSM consists of two separate tools: one is a multimedia presentation authoring tool and the other is a virtual collaboration tool.

Collaboration among experts, such as medical doctors, should be conducted in such a way that all the participants should have prior knowledge about the case that they are about to discuss before the discussion begins.

There has been some work on multimedia authoring such as Bulterman et al.[6]. Their approach is to partition the concerns for authoring environment. The proposed authoring tool offers easy-to-use and efficient multimedia presentation authoring that utilizes Synchronous Multimedia

Manuscript received November 4, 2010; revised November 14, 2010; accepted November 24, 2010.

Chuleui Hong is with the Department of Computer Science, Sangmyung University, Seoul, 110-743, Korea (Email: hongch@smu.ac.kr)

Integration Language (SMIL) [7,8]. It helps the meeting initiator present the case that needs collaboration by providing the capability of synchronizing video and power point file that includes medical images. The uniqueness about the VCSM is the connection between synchronized multimedia presentation authoring and real-time digital image annotation.

In virtual collaboration, annotations use layered architecture that separate images from annotation. In the annotation layer, various symbols such as arrows and/or polygons those are associated with x-y coordinates that represent the location on the image. Those symbols are associated with valuable information such as experts' opinions on the area of interest in some images. All the discussion may be recorded in real-time using XML, as proposed by King et al. [9], and the use of web-based database for permanent storage. With the proposed virtual collaboration tool, users can place a symbol on the location of interest in a digital image and associate it with their opinions and multicast those to all the participants in the discussion so that they may be able to exchange their opinions.

Each submitted text opinion along with their coordinates is saved into an XML file during the discussion for fast loading and retrieval. When the discussion is over, all the opinions that are saved in an XML file may be saved into web-based database for future context based intelligent search. Location (coordinate) information can be used to retrieve opinions on some particular area of the images when selecting the region by polygon, which is typically a convex polygon, and use it as queries to the database for the location and associated opinions.

II. DICOM STANDARDS

The proposed virtual collaboration system will allow researchers or experts (primarily medical doctors) to collaborate in providing quality healthcare. Through this architecture they can visualize Digital Imaging and Communications in Medicine (DICOM) [10] compliant medical images, de facto standard for medical digital images, and exchanges ideas and opinions. The whole virtual collaboration sessions with opinions and image annotations from participants are stored into an XML file as part of the patient demographics information and can be retrieved later.

The DICOM standard aims at providing a protocol for transmission of digital images, and other information, with in the field of medicine. Chiefly concerned with diagnostic medical imaging, DICOM aims to allow for universal transmission of medical images collected from disparate technologies ranging from photographs to x-rays to MRIs. The explicit intent of this standard is to facilitate Picture Archiving and Communications Systems (PACS) applications, which will provide a uniform access to a growing body of medical information as well as communication between involved technologies.

Since DICOM allows for standard transmittal of medical information, with a focus on diagnostic images, its use on a virtual collaboration among doctors should be obvious. Indeed, such a project as this, which seeks to warehouse images and supplemental dialogues, falls right in line with the type of project this standard was meant to support.

The success of virtual collaboration depends upon its access to the full breadth of medical images and information available. First of all, this application must be well-integrated within the system of medical imaging so that uploading of freshly produced images can processed quickly and conveniently, if not directly, from the medical technology to the collaborative technology. Through support and use of the DICOM standard, it will be more likely those doctors using the virtual collaboration will be able to make wide searches of world-wide databases and archives of medical images. Conformance to the DICOM standard will also facilitate distribution of information created through actual virtual collaborations back into the medical community.

The creation of the DICOM standard began with the following goals:

- Facilitate the development and expansion of picture archiving and communication systems that can also interface with other systems of hospital information.
- Allow the creation of diagnostic information data bases that can be interrogated by a wide variety of devices distributed geographically.

This project embodies these goals of the DICOM standard.

III. MULTIMEDIA PRESENTATION

3.1 The necessity of synchronization and SMIL

By far the most effective way of communication is face-to-face meeting. In virtual collaboration community, resembling it as closely as possible would be the eventual goal. In order to achieve this goal, video and other multimedia elements become inevitable choices for the communication.

In this section, we would like to consider multimedia presentation and associated challenges. When used with PowerPoint file, video must be synchronized with each slide so that the audience knows what the video is referred to. The following things must be considered when synchronizing multimedia elements:

1. Duration of each slide (how long should it be displayed)
2. Position of each multimedia element such as video, slide, and/or digital image on the screen
3. Order of each element (which one shows up first)
4. Types of supporting multimedia file and its player

Multimedia presentation is especially useful when there are digital images or textual descriptions that need visual explanation. Video needs to have appropriate position and size and each slide needs to have appropriate duration so that during the execution of video corresponding slides would be displayed. Depending on the connection speed, appropriate bit-rate should also be selected. The whole process is called multimedia synchronization. For such purpose, there is W3C recommendation [11] that describes Synchronous Multimedia Integration Language (SMIL), which describes multimedia presentations. SMIL became an official recommendation in August 2001. It is a collection of XML elements and attributes that can be used to describe the temporal and spatial coordination of one or more media objects. With SMIL, different media objects can be combined into a single coherent multimedia presentation. SMIL can be used to manipulate location, size, type of media objects, bit-rate, etc.

3.2 Multimedia presentation authoring tool

There have been many approaches in adding multimedia to presentations, but many of them are fairly complicated and require expensive equipment. In our approach, we are proposing an easy and inexpensive and yet efficient way of adding multimedia to presentations using SMIL.

The concept was implemented using Java technologies such as JMF[12] and JDOM[13]. It captures video and audio from a web cam and creates multimedia presentation by merging PowerPoint presentation slides. One of the hard parts in creating such material is to specify synchronization relationships among media elements. In our model, we synchronize the video and audio with the PowerPoint slides by creating a SMIL file that specifies the relationships among the elements in the presentation. A SMIL file can specify the duration of each slide, location and size of video and slide and so on. Once the presentation file is created, the proposed authoring tool can edit existing presentations.

Main modules are described as in Fig. 1:

- The Initialization Module initializes a user session. It creates or opens an existing presentation based on user selection and initializes other modules in the tool.
- The JMF video capture module records video and renders it to the screen for monitoring. Its video playback module can play an existing video selected by user.
- Timer is triggered by user command. When the timer is started, it starts video capturing module and marks the beginning of the presentation displaying the 1st slide. As user moves on to next slide, it invokes the slide loader and keeps track of the time duration for each slide for synchronization with associated video. When the timer is stopped, it stops the video capturing and marks the end of the presentation.
- Encoder compress the size of AVI-formatted video and transforms it for Real Player™
- The SMIL Generator compiles the information about slides and time stamp for each slide, captured or pre-recorded video and synchronizes those into one SMIL file that enables synchronized multimedia presentation.
- The SMIL Editor is used for fine-tuning the synchronized presentation. The editor module reads the time stamps from the generated SMIL file. It can add or delete part of slides and videos and change the duration of slides. It also inserts new videos into or deletes the some portions of middle of existing video. It also ensures that the accumulated duration of the entire slides is the same as the video running time.

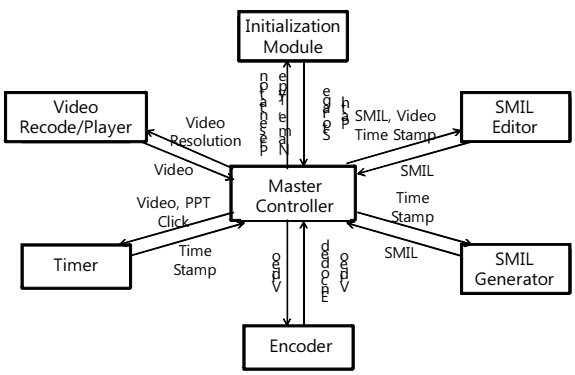

Fig. 1. Block diagram of multimedia presentation tool.

The resulting presentation is in SMIL format, which means the following:

- Media is in Real Player format – encoded video synchronized with slides and other media elements, if any
- A SMIL file defines relationships among the media elements used in the presentation
- The SMIL file is playable by Real Player™ in our approach

Since SMIL is based on XML, an element and its attribute are identified and wrapped by tags. The elements and attributes in the generated SMIL file are consisted of the following:

- Region for video and PowerPoint slides: Regions are assigned to video and PowerPoint slides in each SMIL file using the element <*region*>. Each region is assigned a unique SMIL-ID.
- SMIL-ID: An ID is assigned to each child element of a SMIL element e.g. video or <*img*> in the example code.
- Parallel execution of child elements: Synchronized video and PowerPoint that are executed in parallel are child elements of <*par*> tag.
- Duration time: The <*dur*> attribute defines the duration time for each PowerPoint slide.

The example in Fig. 2 is a SMIL file in action and corresponding source code. The markups for each region and media declarations are clearly shown. SMIL files contain only the references to media objects, not the actual media objects itself. Some SMIL players are currently available [14,15]. The SMIL player we chose was RealPlayer™ [16] due to its popularity.

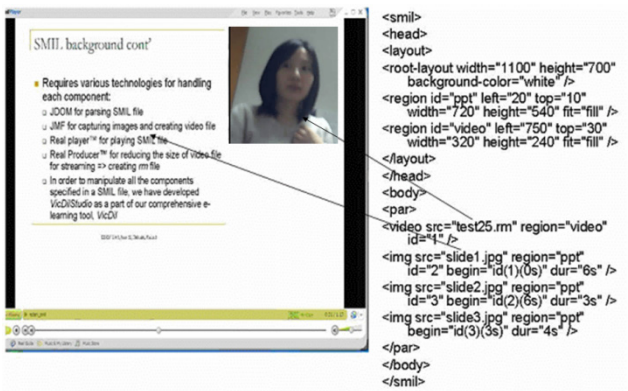

Fig. 2. A multimedia presentation and its corresponding SMIL source file.

IV. DISCUSSION AND OPINION RECORDING USING A VIRTUAL COLLABORATION TOOL

Even though face-to-face meeting is the best way for collaboration, doing so may not be always possible due to geographical and time constraints. In such situation, Internet based applications may be of good help such as email, instant messenger, chatting, IP-based audio or video. With such methods, virtual collaboration may be carried out but the Internet-based collaboration has its own limitations, e.g., discussions that involve images and location because in cyber space it is difficult to locate certain position precisely and dynamically.

To help resolve such issue, the proposed collaboration system is an Internet based distributed system that enables Internet-based recordable discussions over the images. It is a socket/servlet based three-tier distributed system as in Fig. 3.

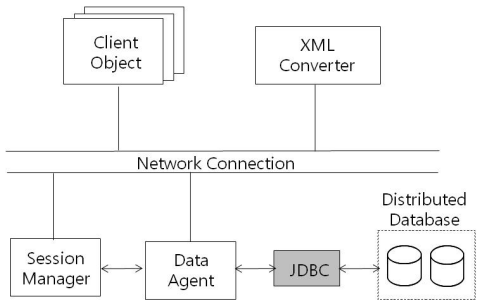

Fig. 3. Three-tiered architecture.

The first tier is made up of client objects. Client objects are Java applets. They communicate with Session Manager or SM, which is the middle tier of the system. SM receives connection requests, participants' opinions and associated x-y coordinates from the front tier and fulfills the requests by invoking XML converter. The SM multicasts the received information to the first tier, which is client objects that are in its active connection queue so that all the client objects can be synchronized. Active connections of the client objects can be periodically examined by SM by sending out query packets to the clients and verifies their reply. Only those who replied to the query will remain in the active queue. Finally, the third tier is the distributed databases such as SQL Server or Oracle.

When the client objects and SM communicate, parameters can be passed as objects that contain participants' opinion or request by way of serialization. On the server side, SM invokes XML converter and have the objects saved into an XML file so that it can be available for the client objects for searching and fast loading. XML converter is a distributed object that converts an object, which contains participants' opinions, into an XML file. When SM polls data from the database, XML converter converts text or any other types of information into XML files for the aforementioned reasons.

4.1 Motivation

When digital images are used in the discussion, it is necessary to point out where the points of interest are. In virtual collaboration, since participants cannot see each other, it is important to know where the point of interest on the image is and what the associated opinions are. In general, digital marking may be possible by drawing some type of symbols on the image. However that may damage the original image

and represent only the limited amount of information. The following things need to be addressed when annotating digital images:

1. Separation of images and labels or opinions: In order for the annotations to be effective, labels must be separated from the images so that the original images will remain untouched and allow as many annotations as needed for the image. It will help manage annotations independently.
2. Association of symbols with various types of information on to the location of interest in the image: the symbols should have some meanings so each symbol need to have some association with various types of information such as text, image, voice and video along with location information
3. Search opinions by location in the image or by keyword: Once annotation for the image is done, it should be able to search either by location or by keyword. When searched by location, x-y coordinates that are limited by the size of the image need to be used to select the locations.
4. In case of real-time collaboration, such information should be delivered to the entire participants in real-time.

We propose a virtual collaboration tool with aforementioned considerations in mind.

4.2 Requirements of virtual collaboration

Discussions among medical doctors require various types of medical images, such as X-ray, MR, CR, CT, SPECT, and Ultrasound. In addition to that, lab results and patients' records are necessary to determine the status of the patient. Since patients' information is considered private, security of the patient information should be one of major considerations. Improper usage of patient information often ended up with lawsuits or some other difficult situations.

- **Medical data**

All the necessary medical records for the collaboration should be collected based on its entirety. Partial information may lead to wrong conclusion and could cause tragic consequences. Sources of the information vary from radiology department to patient record room. Scanners, digital camera, and digital camcorder can also be used to capture medical images. All the information should be in digitized form before the collaboration begins. Medical images are meaningful only when patients' records are provided. Without patient information, it may not be proper for the doctors to determine what the symptoms are. Due to the sensitivity of medical data, DICOM files should only be used in secured environment.

Fig. 4. DICOM data display.

- **Security**

All the participants in the discussions that use patient information have to go through authentication process in order to prevent unauthorized access to the information. Only the authorized participants should gain access to the discussion. It is also necessary to check the active connection of established participants' connection regularly so that the host of the discussion can have actual participants on its list.

- **Discussion**

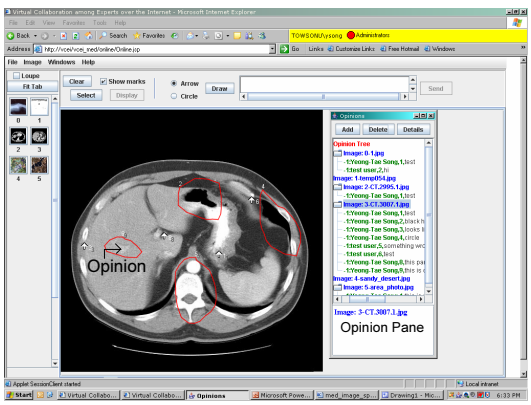

Fig. 5. Opinions shown as opinion pane and associated symbols.

Active communication among the participants will be performed during the discussion session. All the relevant image files should be available to the entire participant throughout the discussion. Each participant's opinion will be recorded and be available for the future reference. When a participant discuss about certain area of an image, a polygon or an arrow will be chosen and placed at a point of interest. Then the opinion from the participant will be associated with the symbol and then displayed in the opinion pane as in the Fig. 5. An XML file will be used to record the symbols and opinions.

- **Search Capability**

The virtual collaboration system is for the experts in their field and their opinions are worthwhile to record for the future reference. All the discussions should be collected and stored in such a way that the information can be retrieved by the keywords, by the type of symptoms or by the certain area of interest in some medical images. When searching for experts' opinions in some area of interest, participants or users simply draw a polygon in the area of interest. Then the symbols located in the area will be selected and displayed in tabular format shown in the Fig. 6.

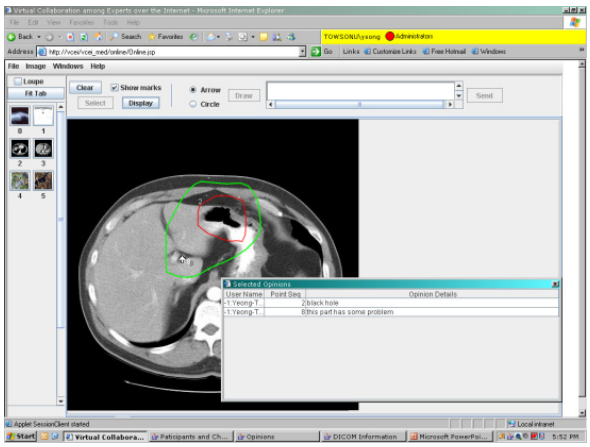

Fig. 6. Search result shown as a table by selecting symbols on the image.

V. APPLICATIONS USING VIRTUAL COLLABORATION SYSTEM FOR MEDICINE

Virtual collaboration among the medical doctors can be performed by the following manner:

1. A doctor comes up with a case that needs discussion with other doctors who have the specialty in that field.
2. The doctor prepares multimedia presentation regarding the case using the proposed presentation tool. The presentation uses digitized medical images and other available resources such as patient information and/or lab results, if any. The images used in the presentation will be available for discussion. When finished, they will be uploaded to the designated cyber conference room.
3. The doctor then specifies the time and date and asks potential participant(s) to log on to the designated web site where the multimedia presentation for the case will be available for them to look at before the discussion begins. Each participant is requested to watch the presentation before the specified time.
4. On the specified date and time, participant(s) log on to the web-site and the doctor who requested the meeting becomes a host of the meeting. Discussions can begin immediately as in Fig. 7.
5. Participants express their opinions through their web browser along with the information on area of interest in some digital images presented. Some relevant digitized material such as medical images can be uploaded from their location, if necessary. Their opinions are saved into an XML file and eventually saved into web-based database such as SQL server.
6. Discussion session continues and opinions exchanged among the participants until they reach some conclusions about the case. Chat window is used for all other messages that are not directly related to the case. The messages used in the chat window will not be recorded.
7. When finished, the host can call the meeting off using chat window.

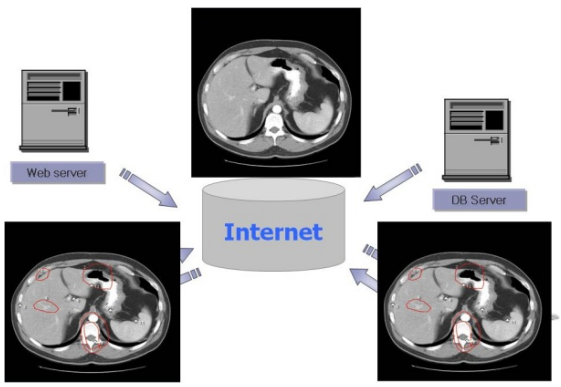

Fig. 7. VCSM in action.

The VCSM can be used to get second opinion from specialists. In this case, the system will be used asynchronously. The steps are similar to the previous case but there is no need for all the participants to be online at the same time. The requester of the second opinion needs to prepare all the images and other supporting document regarding the case. Once they are prepared, they will be uploaded for the specialists to see. The specialists then examine all the uploaded information and post their opinions using symbols over the images and other supporting documents regarding the requested case.

Similar to the second opinion case, it can be used in e-learning to training the medical doctors. Experts on the images can post all the symbols and their associated explanations and hide the symbols from the trainee. The trainees then post their own opinion for the images and compare with that of experts'.

VI. CONCLUSIONS

We have proposed a web-based virtual collaboration and multimedia presentation authoring system for medicine (VCSM). Our approach makes use of spatial database with medical images so that the participants in virtual collaboration may be able exchange opinions using spatial data associated with medical images.

We have provided two different modes for collaboration, synchronous mode and asynchronous mode. Due to the characteristics of medical data which contain many digital images, we have introduced multimedia presentation authoring tool that enables multimedia presentation over the Internet. The tool synchronizes video and PowerPoint file and encodes it for streaming. All potential participants in a discussion are required to watch multimedia presentation from their browser before the discussion begins.

Discussions were performed using layered structure that enables the separation between digital images and participants' opinions. During the discussion, experts' opinions can be collected and saved for future reference. The posted opinions have association with some kind of symbols, arrow or polygon that represents locations on the images so that it can be searched by location or by keywords.

REFERENCES

[1] W. Choi, H. Kim and M. Oh, "Video Conference System Construction that Apply Real Time Video Encryption Techniques", The Journal of KIMICS, Vol. 6, No. 6, 2002
[2] D. Roberts, T. Duckworth, and et al., "Comparing the End to End Latency of an Immersive Collaborative Environment and a Video Conference", Proceedings of IEEE/ACM International Symposium on Distributed Simulation and Real Time Applications, 2009
[3] Gary Olson, Stephanie Teasley, Matthew Mietz, Merrick Cogburn, "Collaboratories to Support Distributed Science: The example of International HIV/AIDS Research", Proceedings of SAICSIT 2002, pp 44-51
[4] Siriwan Suebnukarn, Peter Haddawy, "A Collaborative Intelligent Tutoring System for Medical Problem-based Learning", Proceedings or IUI'04, Madeira, Portugal, January, 2004
[5] M.. Pourdadash M.. Hamidreza and et al. "A Web-Based Framework for Distributed Medical Image Processing Using Image Markup Language (IML)", Proceedings of UKSim European Symposium on Computer Modeling and Simulation, 2009
[6] H. Pooshfam, M. Rajeswari, and D. Ramachandram "A web-based framework for resource sharing and collaborative medical image annotation", Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia, 2009
[7] Dick Bulterman, Lynda Hardman, "Structured Multimedia Authoring", ACM transactions on Multimedia Computing, Communications and Applications, Vol. 1, No. 1, February 2005
[8] SMIL Tutorial, http://www.w3.org/TR/SMIL2/
[9] C. Kim, D. Lim and H. Chung, "SMIL Authoring System for Multimedia Object Presentation", The Journal of KIMICS, Vol. 13, No. 7, 2009
[10] Peter King, Patrick Schmitz, Simon Thompson, "Behavioral Reactivity and Real Time Programming in XML", DocEng '04, October, 2004
[11] DICOM Standard Status, http://www.dclunie.com/dicom-status/status.html
[12] W3C Recommendation http://www.w3.org/TR/2005/REC-SMIL2-20050107/
[13] Java Media Framework JMF http://java.sun.com/ products/ java-media/ jmf/ index.jsp
[14] JDOM http://www.jdom.org/
[15] GRiNS Player for SMIL 2.0 http://www.oratrix.com/
[16] X-Smiles SMIL 2.0 Player, http://www.xsmiles.org/xsmiles_smil.html
[17] Real Networks, http://www.realnetworks.com/products/free_trial.html

ACKNOWLEDGMENT

This work was supported by the Sangmyung University Research Grant of 2010.

Chuleui Hong received B.S. degree in Hanyang University, Seoul, Korea in 1985, and his M.S and Ph.D. degrees in Computer Science at New Jersey Institute of Technology, USA and University of Missouri-Rolla, USA in 1989 and 1992, respectively. He was a senior researcher in Electronic and Telecommunications Research Institute, Korea, from 1992 to 1997. He joined Sangmyung University as a faculty member in 1997 and is currently professor of Computer Science Department at Sangmyung University, Seoul, Korea. His research interests include Parallel and Distributed System, Optimization Algorithm, multimedia Application, and Intelligent Agent.