JoinSee: A Real-Time and Collaborative Hyper-Media System for Participatory Performances in the Opera of Meaning

Shuichi KURABAYASHI a, Shlomo DUBNOV b and Yasushi KIYOKI a

a Graduate School of Media and Governance, Keio University
Endoh 5322, Fujisawa, Kanagawa, Japan

b Music and Computing in the Arts, University of California, San Diego
9500 Gilman Dr. La Jolla, CA 9209 USA

Abstract. In this paper, we propose a real-time and collaborative hyper-media system that introduces database-enhanced collaborative models and multimedia processing models for creating improvised performances in the Opera of Meaning. The system provides a main story media and a corresponding shared canvas that is shared among Internet-wide user communities. Our shared canvas mechanisms make it possible to describe and share users' ideas and impressions about the main story media. The key technology of this system is a timeline-dependent and script-driven live performance engine, which provides users with ECA rules to express and characterize the users' ideas and impressions by using existing multimedia data such as video files and image files. The system provides directors and participants of improvised performance with a set of database operators for controlling and contributing to the performance. The system motivates users to contribute to the performance by exploiting users' own media libraries and existing web services. We have implemented the prototype system which is applicable to the existing video and image files on the Web.

Keywords. Collaborative System, Multidatabase Systems, Active Database Systems

1. Introduction

The popularity of web conferencing and content sharing has been driven by a desire to increase productivity with a way to meet instantly and operate cost-effectively on digital media and electronic text from distant locations. Collaboration through remote web presence also becomes increasingly prevalent in entertainment situations such as telematic performance, multiuser computer games and participatory film and music performance events. While most of the systems try to improve users experience by increasing the fidelity and bandwidth of the user videos, sound or graphics, an important aspect of user engagement related to planning of the joint activity is usually left out of the system design, and is determined by the personal skills of the presenter and the interest of the attendees. Increasing the level and quality of user representation through video or avatars does not solve the problem of lack of presence, since there is no way to assure that the person at the other end of the communication line is actually committed to participate in the event. This difficulty in creating a suspense of disbelief
about a remote event seems to point at the importance of creating an intensive social interaction that is distributed among the attendees and yet coordinated around a common goal or theme.

In this paper, we present JoinSee that is a real-time and collaborative hyper-media system which allows implementation of different multi-user intelligent applications for discussion and content sharing. In JoinSee we propose a system and a model of participation that is based on a shared multi-user canvas, with an underlying model of knowledge described by a script and a time-line, related to a main movie, story or central content that is multicast to the attendees. The shared canvas provides a "back channel" for user communication that happens in parallel to watching the main story. Using semantic database management, additional contents that are relevant to the main story are made available to the users. These contents can be shared at predetermined times using the shared canvas mechanism. The JoinSee system tries to address the important issue of sharing a common virtual space by establishing a conceptual semantic and impression data environment that is shared among multiple users in a distributed manner. This setup allows mixed attendance modes that combine first person interaction among attendees directly or through a moderator, and a third person mode of viewing with regards to the main story, with its associated freedom of imagination and deliberation. This mixed operational mode has both advantages and disadvantages. As shown in Fig.1, JoinSee provides several canvases for users connected via Internet. There are two types of canvas: a public canvas which is visible to all users, and a private canvas which is dedicated to each user.

![Fig. 1 A Basic System Overview of JoinSee](image)

The main advantage is creating an environment for social interaction and possibility for active participation that is beyond passive third person viewing of the main presentation or limited occasions of one to one presenter attendee interaction. On the other hand, the multi-user interaction in parallel with the main presentation demands multi-tasking and split attention viewing, which might be distracting or more difficult to the users. Realizing “Opera of Meaning” [1] environments in the Internet requires several new techniques to be developed. In order to apply a context-dependent multimedia search engine to live performance, a balance between precision and real-time needs must be achieved.

The key technology to realize JoinSee is an active database mechanism for meta-level multidatabase environments to promote dynamic information sharing on the
public and private canvas. The meta-level multidatabase system [4][5] is an abstract and higher level layer beyond local databases, and it would be constructed independently to the local database systems to connect and integrate those databases. We have developed the meta-level multidatabase system that offers active database functionalities for the multimedia contents on the Web [6][7]. The meta-level active multidatabase system supports Event-Condition-Action (ECA) rule mechanisms. ECA provides a general and powerful syntax and semantics that alleviates the programmers from the implementation of monitoring a large size of dynamic data [9][10][11]. ECA rule specifies when the rule is triggered (Event), what state to be monitored and detected (Condition), and how to perform database operations if the condition is satisfied (Action). JoinSee is an application of the meta-level multidatabase system active database functionalities. In JoinSee we have developed timeline control mechanisms in the active multidatabase system in order to supports a timeline-dependent live performance on the public canvas.

This system offers new ways of connected collaboration, which are not possible with existing communication-oriented tools such as VoIP, Pub/Sub systems [2][3], and video hosting services. The contributions of this work are the following:

- JoinSee provides a shared canvas system that consists of both a media player and a corresponding free-space. The media player is used for displaying a continuous media data, such as video data, as main story. The free-space is for displaying images and web snippets. The shared canvas makes it possible to describe and share users' ideas and impressions about the main story media. JoinSee provides a director and participants of improvised performance with a set of database operators for controlling and contributing to the performance. A current access control mechanism in the shared canvas is a straightforward endeavor; however it is very important to inspect emotional relevance between posted objects and the context of main story.

- We present a timeline-dependent and script-driven live performance engine for the shared canvas. The live performance engine integrates existing databases according to the dynamic context of main story media. The live performance engine facilitate to express and characterize the users' ideas and impressions by using existing multimedia data such as video files and image files. The search can be performed using third-party search services like Flickr, YouTube or Google on legacy databases, or on local closed-world database, and combination of the two.

- Third, but not least, we show a practical implementation that offers the shared canvas on the web. The prototype system integrates existing multimedia data along with timeline of the main story media. Our implementation contains a modular framework for integrating actual multimedia data including H.264 movies, JPEG images, and web search results. Moreover, we show an implementation method to utilize Service-Oriented-Architecture (SOA) technologies for integrating existing web services into JoinSee system.

2. JoinSee System Architecture

In the Opera of Meaning environment, there are two types of users: directors who manage the entire contents, timing, and quality of a performance, and participants who watch the performance. JoinSee renders the performance on the shared canvas that can be watched from distributed clients. The shared canvas is a unique service by JoinSee that allows sharing of user generated contents, including results of multidatabase search
between users. The shared canvas concept consists of two elements: private and public. The private canvas is like a sketch pad where individual users can browse various data using their own machines. The public canvas presents only part of information from individual users that are considered most relevant, either by posting of the users to the shared canvas or by filtering of the user browsing actions by a moderator to show only more meaningful elements or solicit specific actions. The JoinSee system provides a set of database operators to directors and participants for improvised performances in the Opera of Meaning environment.

Fig. 2 shows a screenshot of the Internet-Wide Opera of Meaning environment realized by JoinSee. There are two main components: 1) the main video that shows the central story for discussions, 2) the public canvas that is synchronized to the main story video. The system provides synchronization mechanisms that control the state of media-player of all participants. Contents on the public canvas are changing along with timeline. Users upload various multimedia data to the system. The system evaluates the relevance between the uploaded media data and the main story in order to display the relevant information on the public canvas.

### 2.1. Data Model

The JoinSee server system manages the shared canvas and its timeline associated to the main media story. Our data model consists of three elements: a main story, a timeline, and tracks. A main story is a continuous media data such as video data and music data. The main story is a central object of the performance, where every user watches the main story and posts the relevant multimedia object to the shared canvas. A timeline represents a sequence of events associated to the time period of main story.
The shared canvas consists of multiple “tracks” as shown in Fig. 3. A track is a data structure to specify a set of media objects along with timeline. Each track is associated to the one participant. The system provides $i$-tuples tracks for $i$ users. When a user posts media objects to the shared canvas, the objects are inserted into the corresponding track.

We denote a shared canvas by $s_i \in S$, where $S$ is the all main story media installed in the system. We denote a set of tracks that belong to the main story $s_i$ by $T_i$. We denote a track associated to a user $u_i$ by $t_{ij}$. $t_{ij}[n]$ denotes a $n$-th media data object belonging to the track $t_{ij}$. A media data object $t_{ij}[n]$ is a three-tuple: $t_{ij}[n] = \{\text{URI}, \text{MIME-TYPE}, w, h, m\}$, where URI is a global identifier of original data sources. MIME-TYPE denotes a media type of media data, $w$ and $h$ denote a bounding box of media data which is to be rendered, and $m$ is a metadata description that is represented with a set of key/value pairs $<f_0, v_0>, <f_1, v_1>, \ldots, <f_k, v_k>$ with $f[0\ldots k]$ being a feature name and corresponding $v[0\ldots k]$ being a value.

2.2. Automatic Video Analysis Method for Creating Metadata of Main Story

Fig. 4 shows a color-impression feature space that represents color-based dynamic contents such as image and video data. We have designed the schema for the metadata structure for representing color-impression of image and video data using 120 chromatic colors and 10 monochrome colors defined “Color Image Scale” [8] based on the Munsell’s HSV color system. We used 183 words, which are also defined as cognitive scales of colors in the Color Image Scale, as metadata feature set for tracks. Fig. 5 shows the timeline-dependent color-impression analysis for video data.
features. By correlation calculations between 183 color schemas and 130 basic colors, this system extracts the color-impression for each frame, and creates a sequence of color-impressions of the video along the timeline as shown in Fig. 6. This system offers new ways of querying, which are not possible with few words, by exploiting existing video data libraries. We have implemented the prototype system that is applicable to perform queries over the various video files on the Web.

Fig. 6  Visualizing Impression-Transition in Video Data by Using 183 Impressive Color Schemas

3. JoinSee Server System and Shared Canvas Mechanisms

In this Section, we present an implementation method of JoinSee and its shared canvas mechanisms. Fig. 7 shows that the JoinSee server system provides the following mechanisms for the shared canvas: 1) cross-media association mechanisms, 2) privilege control mechanisms for the public canvas, and 3) synchronized media player. The JoinSee system provides semantic privilege control mechanisms that evaluate relevance between each user’s postings to the shared canvas and the context of a main story. The system applies automatic metadata generation functions to the posted objects for inspecting timeline-related relevance to a main story media.

JoinSee consists of both the central server system and several client systems. The JoinSee server system uses the SOAP specification as communication protocol between components. The client systems provide functionalities for rendering multimedia data. The server system is built on the active multidatabase system, and the server system is managed by ECA rule language. A director of the performance defines ECA rule to integrate multimedia data into the shared canvas according to contexts in the improvised performance. One of the advantages to introduce ECA rules is that an ECA rule automatically reacts to users’ behavior such as posting, seeing, taking, and searching. The server system generates multimedia contents dynamically and multicasts the generated contents to each client. Those I/O communications are driven by ECA rules which are defined by the director of the performance. JoinSee supports a look-ahead mechanism for multimedia objects on the shared canvas by applying ECA
rules in advance. By applying this system architecture, we incorporate a wide variety of legacy information systems including modern web services, such as Google, Flickr, and YouTube, into the collaborative media creation environment without modification of the system. In addition, we have implemented fundamental social networking mechanisms for JoinSee. The social networking aspects of JoinSee are central for creating the user experience and social impact of the system. People want to watch together movies and also link their own experiences to parts of the movie. This means that people who upload or link their own images, text or graphics to different parts of the main movie can “push” or solicit these contents to their friends.

3.1. Collaborative Media Creation

The all reactive behaviors provided in JoinSee are implemented by using ECA rules. The JoinSee system realizes collaborative media creation by performing the following five steps:
Step-1: Analyzing a main story media data: The JoinSee server applies the color-impression analysis method to the video data to generate metadata.
Step-2: Preparing media data: The JoinSee server integrates existing multimedia database for preparing a set of candidate data to be displayed in the shared canvas.
Step-3: Starting the performance: The system monitors the state of every private canvas. Users search the appropriate media data by using web services, and they post the find media data into the private canvas. When the system detects a change in the private canvases, the system invokes corresponding ECA rules.
Step-4: Evaluating the private canvas: The invoked ECA rule applies automatic metadata analysis method to every media data in the private canvas. And then, the
system computes relevance between the media data and the main story. When the relevance score is over a threshold, the system registers the media data in the public canvas.

Step-5: Multicasting the public canvas: When the system detects a change in the public canvases, the system starts sending the contents of public canvas to each client.

4. Conclusion

In this paper, we have proposed the real-time and collaborative hyper-media system that introduces database-enhanced collaborative models and multimedia processing models for creating improvised performances in the Opera of Meaning. A unique feature of our system is that it offers the active database mechanism for meta-level multidatabase environments to promote dynamic information sharing on the public and private canvas. As a future work, we are planning to expand the scalability of the system for offering the collaborative environment to a bunch of remote users.

References