# Table of Contents

The Issues Regarding the Distance between the Surface and the Displayed Objects  
Minoru Kobayashi  
1

Collaborative Creation of Interaction Patterns for the Use in User Interface Generation  
Benjamin Weyers, Johannes Nanninga and Wolfram Luther  
7

Co-Musicator: A Web-Based System for Musicians to Support Recording, Playing, Indexing and Discussing  
Takehiko Sakamoto, Shin Takahashi and Jiro Tanaka  
15

A Networked Discussion Tool for Shogi  
Yoshikazu Tagashira, Hiroyuki Tarumi, Shinji Toshima, Shinji Ohara, Toshiki Kinuhata and Toshihiro Hayashi  
23

A Proposal of Digital Photo Frame Sharing Photos and Comments  
Satoshi Ichimura  
31

Learning Log Navigator: Supporting Authentic Learning Using Ubiquitous Learning Logs  
Kousuke Mouri, Hiroaki Ogata and Songran Liu  
39
Issues Regarding The Distance Between The Screen Surface and Displayed Objects

Minoru Kobayashi
Meiji University, 4-21-1 Nakano, Nakano-ku, Tokyo 164-8525 Japan

Abstract. This paper refers to two research prototypes in discussing how the perceived distance impacts the understanding of objects displayed by visual communication media. Gaze awareness of ClearBoard is found to be degraded as display size falls, and perceived object size becomes indefinite in the absence of distance cues. Based on these results, design issues for the next generation of remote collaboration environments are discussed.

1 Introduction

The display screen works as a window between the space of our physical world and the space of the media world. For example, when showing a photograph, it is a window to the space of the scenery; when used in a video conferencing system, it is a window between local and remote sites. We use visual cues provided by the screen to better understand the structure of the space beyond the screen, and judge the location or size of objects in the image.

Some high-end video conferencing systems or telepresence systems[1] provide users with sense of being together by designing the environment so that the remote parties occupy identical spatial structures. Such systems encourage the users to feel that the remote site is an extension of their space, and can judge the distance to other participants or guess the meaning of their facial directions or gestures, which can improve the communication quality.

This paper aims to provide clues to start a discussion on designing visual media that makes it easy to understand the spatial structure of the space beyond the screen, which would result in a better understanding of the displayed objects. We start by discussing how the perception of distance affects the understanding of displayed images, based on the author’s experience in designing two visual communication systems, one supports gaze awareness, and the other supports the sharing of perceived size.

2 Gaze Awareness in ClearBoard

When we discuss documents in the real world, the space for conversation and the space for document sharing occupy a single contiguous space. However, in many teleconferencing systems, they are divided into windows for conversation and
windows for shared documents. ClearBoard[2] is a remote collaboration medium that integrates these spaces in a natural manner, so that the users can use their communication skills, such as gestures or facial expressions, as they would in daily life.

To achieve the integration, ClearBoard employs the metaphor of “transparent glass board.” That is, the display surface of the telecommunication terminal is designed to replicate a glass board placed between the user and the partner, on which they can draw together and see and talk to others as if they are talking thorough and drawing on a transparent glass board.

In ClearBoard, thanks to the half-silvered mirror covering the display surface, a single camera above the display screen captures the user’s face and body, as well as the shared drawings and gestures made on the screen. The captured image is virtually equivalent to the image captured through the screen, and thus the metaphor of “transparent glass board” is realized by ordinary display devices. Users of the ClearBoard see the drawings on the screen, see the partner behind the screen, just as they would see when standing on their side of a transparent glass board.

The original goal of ClearBoard was to convey the spatial relation between the drawing surface and the user’s body, so that users can use gestures or facial expressions to enhance the other’s understanding of the drawings. While developing the system, we found that it actually exhibited a more important feature, the so called “gaze awareness;” this allows it to convey gaze direction more effectively than ordinary whiteboards or drawing tabletop surfaces. Gaze awareness is the ability to monitor the direction of the partner’s gaze and thus his or her focus of attention, and ClearBoard supports precise gaze awareness.

In communication, we read the partner’s gaze to understand their focus of interest, or to foresee their action. We also use gaze as expressions; we focus on the partner’s face or their drawings to show how attentively we are listening, or when we find something interesting, we shift our gaze to that object to induce a shift in conversation topic. ClearBoard establishes a shared workspace between partners and ourselves that allows us to easily switch our focus between the partner’s face and screen objects, which makes communication more fluid and effective. With ClearBoard-1, the users refer to the partner’s gaze and use it in communication far more often than the tabletop setting[3]. The close pairing of the shared drawing surface and the facial image makes it easier to switch our attention between them, and makes it easier to know what the partner is looking at. This feature of precise gaze awareness is well maintained even in video-mediated systems.

3 Gaze Awareness in a smaller screen system

The original ClearBoard systems were inevitably bulky, because the video projector and the camera had to be placed symmetrical with respect to the screen surface in order to simplify the compensation of the trapezoidal distortion generated by capturing the drawing surface at an angle. In 2003, making maximum
use of digital data processing enhancements of PCs and Internet communication which were continued to advance, I designed a digital ClearBoard, which replaced the optical systems of ClearBoard-1 with digital video processing[4].

This digitalization not only dramatically miniaturized the system, but also yielded flexibility in terms of arranging the camera and display. Unlike the original ClearBoard, which required precise device arrangement, digital ClearBoard well supports a variety of device configurations by setting the digital video processing parameters appropriately. Provided the camera can capture the whole area of the screen, the system can reproduce the drawing surface in its original shape, and display the space beyond the screen as being contiguous with the screen window.

In digital ClearBoard, the cameras do not have to be placed above the screens; they can be placed in front of the screen or even at the side of the screen. It is reported that such camera arrangement changes the image, and affects the impression of the images. Also, the camera distance changes the size of the facial image; when the camera is located close to the screen, the facial image is displayed smaller than that taken at a distance because of the change of the viewing angle needed to maintain the size of the drawing surface. This also affects the impression created.

Even more, the system allows heterogeneous camera arrangements; the cameras do not need to be placed at the same place at both sites, and thus remote partners could see images of different perspective while sharing the same image of the drawing surface. In the paper of 2003, I took that flexibility as a virtue of digital processing; such visual effects could be used to modify the impression. However, from another point of view, this could raise difficulties. Users of digital ClearBoard have difficulty in imagining how close the partner is to the screen, or how large the screen at the remote site is, unlike the original ClearBoard, which set identical terminal devices at both sides and thus made it easy to assume that the partner was occupying the same environment.

By taking advantage of the flexibility of the digital version, I built a 15-inch version of ClearBoard. As there was some concern that gaze awareness might be degraded due to the camera position or image quality, I examined how accurately users could estimate the partner’s target of focus from their gaze. The errors were not always large, the camera arrangement did not significantly impact the preciseness of gaze detection, and the paper concluded that gaze awareness was maintained to some extent in digital ClearBoard.

This paper focuses on the error rate as it may give us clues to understand the users’ perception of the media space. Based on the results shown at the presentation at UIST2003, all errors had the same tendency that is locations were shifted inward to the center of the surface with some variance. It might be the result of that the gaze awareness was “impacted by the sense of distance between the partner and drawings” as described in the paper. In the case of those experiments, viewers seemed to estimate the distances as being shorter than they were in reality. As shown in Fig. 1, when the partner looks at target A, the viewer
Fig. 1. Error of gaze awareness by viewer’s underestimation of distance to the partner may judge the partner is looking at target B, if the viewer underestimates the distance.

One of the possible factors for distance error is the screen size. Though experimental proof was not provided, based on the experiences gathered from the use of the existing ClearBoard system displayed at a public museum, users of ClearBoard-1 seemed to have more precise and easier gaze awareness than was achieved with the smaller screen ClearBoard. The major difference between them is the screen size. ClearBoard-1 has 80-inch screen, which covers the upper body and arms, as well as the partner’s face and hand gestures. The users can see the posture of the partner who is facing the drawing surface. They can see the full extent of hand, arm, body and head, and have enough information to estimate the distance between the partner’s face and the drawing surface. On the other hand, the digital version had a small 15-inch screen\(^1\), which covers only the drawing, a hand, and the user’s face. It is hard for the users to find clues to estimate the distance from the fragments of body parts.

Another possible factor is the flexibility of the system. Unlike ClearBoard-1, the users cannot assume that the partner is directly facing the drawing surface in the same way they themselves do, because they know the images can be adjusted in various ways. These two factors may result in uncertainty as to the distance, and errors in gaze awareness.

It seems reasonable to say that the visual media should provide users with clues to estimate the distance between the screen object and the partner in order to provide precise gaze awareness, though we cannot confirm this without experimental proof. Using large screens, and a common configuration are possible design approaches, but we need some breakthroughs that make it possible to provide those clues without using these approaches that limits the devices, in order to make full use of the various kinds of devices.

\(^1\) Small screen size is not a requirement of the digital version. The small size was chosen because it was thought to be handy in the office environment, and the digital version made it easy to realize.
4 Distance issues in Real Scale Video

In this section, I introduce another instance of a visual communication medium for which there is a need to discuss the distance issue.

Real Scale Video[5] is a project to realize video communication systems that convey an accurate impression of object size. It adds size information to pictures or videos, so that senders and viewers can share the same impression of object size. There were two generations of the system. The first generation system used two-dimensional display, and computed the projection scale to reproduce the object’s size, and display the object at its original size. The second generation system used a three-dimensional display that directly provided a sense of distance, so that users could judge object size based on the visual angle and the distance.

This project started with the finding that even full-scale pictures do not always give the right impression of size. As many factors, such as background color or the direction of the objects, affect size perception, the first generation system employed an adjustment mechanism to account for the user’s bias in size perception and lessen the difference between the full-scale image and the impression. Assuming that some error was due to the error in distance perception, the second generation employed a three-dimensional display, which directly provided depth cues and allowed users to perceive the distance to the objects. As the second-generation system was effective, precise perception of the distance may be important for sharing the impression of size.

In the first, two-dimensional system, the objects displayed at actual size were perceived to be their actual size, if the object was seen to lie on the display surface. However, in many cases, the objects are not seen to lie on the screen due to the composition of the image. For example, objects tend to be perceived as lying on the screen when the camera is directly in front of the screen, but not when the camera is skewed relative to the screen. If it is seen to extend beyond the surface, the object is perceived to be larger than the displayed size on the screen, while if it appears to project in front of the surface, the object is perceived to be smaller than the displayed size (Fig. 2). Based on our daily experience, most objects displayed on TV screens exist beyond the screen (i.e. into the TV-set) because they are normally a part of a scene. Thus objects displayed at actual size on two-dimensional displays may be perceived larger than the actual size.

The impression of size is greatly impacted by the perceived position of the object. It is important to make it possible for users to intuitively understand the structure of the space that includes the screen surface and the objects, and the relations between them.

5 Discussion

This paper explained possible reasons for the influence of the perception of the distance between the display surface and the displayed object on the understanding of the size of displayed objects, by referring to two research prototypes
Depending on the perceived distance from the surface, the impression of object size differs.

of communication media. ClearBoard needs to have intuitive and precise understanding of the distance between the shared drawing surface and the users’ faces to achieve precise gaze awareness. In the Real Scale Video system, the perceived size of displayed objects varied depending on whether the displayed object is perceived to be on the surface or behind or in front of the display surface.

I expect that even with two-dimensional displays, the sense of distance can be conveyed by using techniques that can express spatial cues as to the spatial configuration, such as techniques to include the user’s bodies in the images, or to let the users to know that they share a common spatial arrangement.

Computers and networks are becoming smarter and more ubiquitous, and small devices such as smartphones can be used to join videoconferences. Heterogeneous communication environments in which some participants use smartphones and others use big screen are going to become common, and the distorted spaces created by connecting spaces of different perspectives make it more difficult for users to share the sense of distance or the sense of space. In view of this immanent communication environment, it is important to meet the human interface challenges that must be overcome if we are to create a shared sense of space. I hope the distance issues discussed in this paper will give a clue to such challenges.

References

Collaborative Creation of Interaction Patterns for the Use in User Interface Generation

Benjamin Weyers¹, Johannes Nanninga², and Wolfram Luther²

¹RWTH Aachen University, Virtual Reality Group, Seffenter Weg 23, 52074 Aachen, Germany
weyers@vr.rwth-aachen.de
²University of Duisburg-Essen, Dept. of Computer Science and Applied Cognitive Science, Lotharstr. 65, 47057 Duisburg, Germany
nanninga@gmail.com, luther@inf.uni-due.de

Abstract. Collaboration offers great benefits to the creation of interaction patterns that can be used in a second step, for instance, to generate a graphical user interface semi-automatically. The collaborative aspect in the creation of patterns enables the identification of weaknesses, the solving of errors in pattern descriptions, and finally enhances the quality of the result of this process, such as a use case-dependent user interface implementation. Wikis are a well-established concept for collaborative creation of content. However, these platforms are mainly employed for the description of informal text. The work presented here introduces an extension of the MediaWiki engine adding support for XML-based content. Through the use of XSLT style sheets and processors, this content can be adapted to various application scenarios and prepared for the use in the user interface creation.

Keywords: Wiki, collaborative creation, interaction pattern, user interface generation

1 Introduction

Nowadays, patterns are widely spread in various disciplines of computer science. Especially in software development, patterns are used to create software solutions with a high quality and to reduce costs reusing existing and well established solutions to known problems. The same is true for the use of patterns in the development of (graphical) user interfaces (UIs) by researchers and UI developers.

Nevertheless, the use of UI patterns is largely restricted to the use by a (human) user interface developer because patterns are generally described based on natural language. Alexander (1977) [1] defines a pattern such that “[e]ach pattern describes a problem which occurs over and over again in our environment, and then describes the core of the solution to that problem, in such a way that you can use this solution a million times over, without ever doing it the same way twice”. Thus, the interpretation of a pattern description depends on the developer. Patterns are lightly structured into sections: every pattern has a title, a problem statement, and a description of the solution. Depending on the pattern description concept, further aspects will be defined, such as dependencies...
between patterns or diagrams and examples. This informal nature of pattern descriptions makes it impossible to use them in algorithmic processes, such as in case of the semi-automatic generation of user interfaces.

The creation and reuse of patterns is often hindered by restricted access to pattern descriptions and the creation process of patterns, which is often not open to a broader community. Various pattern repositories exist, such as those published by Tidwell [2], Borchers [3], or Duyne [4], all using different description languages and structures, which further hamper the use of patterns in application development. First efforts to formalize patterns and to use them for the semi-automatic generation of user interfaces have been discussed by Engel et al. [5]. They extend PLML [6], a markup language for the description of patterns, with formalisms, such as ConcurTaskTrees [7] that offer data and information to be fed into a user interface generation framework called PaMGIS. The generation process is further steered by the user who inputs additional information making the whole process semi-automatic.

Based on these findings, we implemented a web-based implementation of a collaborative platform for the development of interaction patterns also aiming at overcoming the problem of a heterogeneous and informal description (format) of patterns. Therefore, we implemented a Wiki extended with an XML processing tool, which is able to process and store XML-based descriptions as well as let the user edit patterns using HTML forms embedded into the Wiki page. The use of Wikis for creating patterns is not entirely new. The first Wiki called “WikiWikiWeb” was established for the community-based description of patterns in object-oriented programming [8]. Still, this Wiki does not use any kind of well-structured document formats or any other type of formalization, as intended by our solution.

As basis for our implementation we use MediaWiki [9] rather than other existing Wiki engines such as UseModWiki [10], because MediaWiki offers a well-established interface for the implementation of extensions and is supported by a broad and lively community. Our implementation comprises an XML transformation extension that is able to load XML-based pattern descriptions and convert them into WikiText that can be interpreted by the MediaWiki engine. WikiText is a text-based markup defining the structure, the format and the content of an HTML-based Wiki page generated by the Wiki engine. The extension is further capable of generating HTML-based forms from the XML-based pattern description, which enables a simple editing of patterns by the user such that she does not have to edit the XML source directly. After the user finished the editing, the extension writes the edited pattern back to the database after it extracted the content out of the HTML form and generated an XML-based representation.

By using PLML, we found an approach for the format-independent and community-based description of patterns. Using techniques such as XSLT and XML Schema, the patterns defined in the Wiki are adaptable to various implementations and above described solutions. Solutions defined in, for instance or XUL [11], can be stored, extracted, and used in concrete application contexts. Also the transformation to extensions of PLML, such as XPLML and PLML v1.2 [12] or the use of ConcurTaskTrees [7] (as discussed above) are covered. Thus, the main contribution of this paper is an implementation of a Wiki (extension) for community-based and collaborative creation of
patterns that is extended by a processing plugin for handling input, storage, and processing of XML. The Wiki is mainly addressed to UI developers and researchers to make UI patterns accessible for the creation of UIs as well as providing a platform for the discussion of UI patterns in the research and development context. The Wiki can be accessed under http://www.pattern-wiki.org. Overall, this approach and the use of the Wiki concept opens up possibilities to achieve collaborative modeling and creation of reusable interaction patterns and use them in semi-automatic generation of user interfaces.

The paper is structured as follows. Section 2 introduces the realized Wiki solution and describes the implemented extensions. Section 3 introduces a possible workflow of usage for the described patterns and identifies ongoing implementation efforts. Section 4 finally gives a conclusion and some future work.

2 PatternWiki – A Wiki for Collaborative Creation of Interaction Patterns

Before starting to highlight the implementation of a Wiki for the collaborative creation of reusable interaction patterns, the following list will summarize the requirements to be resolved by the implementation: the Wiki should be able or enable

- To process XML-based data,
- To apply XSLT style sheets to XML documents from the underlying database to support reusability of patterns in different target languages and platforms,
- The user to edit XML-based pattern descriptions using an HTML form-based editor, and
- To visualize dependencies between patterns contained in the Wiki.

To our knowledge, the use of arbitrary XML formats in Wikis is currently not covered in any Wiki engine implementation making it necessary to implement an extension to an existing engine to reduce the overall implementation costs and profit from well-suited and existing Wiki realizations. Thus, MediaWiki has been selected being best-suited to fulfill the requirements discussed above (see Introduction), especially due to the valuable extension interface provided by it. The implemented extension mainly offers the functionality to process XML documents and translate them into a format the Wiki engine is able to interpret (WikiText of HTML-based forms). The transformation process can be seen in Fig. 1. In the case of querying a pattern from the database through the Wiki internal search engine, the XML source will be delivered together with an XSLT style sheet to an XSLT processor generating WikiText from the original XML source. This WikiText is then delivered to the MediaWiki parser that generates the final HTML page to be sent to the user’s web browser.

Furthermore, the extension has to offer the creation (solid arrows in Fig. 1) and editing (dashed arrows in Fig. 1) of new and existing patterns in a HTML-based form. For that, the extension uses the same XSLT based mechanism creating a HTML form through transforming an empty standard pattern document (in the creation case) or an existing pattern fetched from the database. From the entered data, the new XML file is
directly derived and stored into the underlying database (fine dashed arrows in Fig. 1).

The separation of documents between standard WikiText, HTML forms, and XML documents mainly describing patterns or transformation style sheets is implemented using namespaces. Thus, the resulting Wiki is still using WikiText for articles that do not define patterns, such as the top page or a documentation text for the use of the Wiki, etc. Through the association of data items (Wiki pages) in the database to namespaces, the extension checks in a first step whether the fetched data item is a Wiki page or an XML document. If it is the latter, it further checks whether the document is associated to a schema definition or not. If so, the extension validates the document against the schema and finally passes it to the XSLT processor or stores it if it is valid.

For the Wiki implementation, PLML has been chosen for the description of patterns (see Introduction). Based on the use of XSLT-based transformation, the Wiki extension (by adding it through the Wiki itself to the underlying database) does not mandate this format. Adding a new style sheet to the Wiki using the extension and defining a new namespace, a new format for pattern description is simply added to the existing implementation without the need of writing additional code.

However, PLML is established in the community and represents a finely balanced agreement on pattern elements compared to existing pattern formats and languages. Beside others aspects, PLML defines dependencies between patterns, such as the original pattern definition of Alexander [1] does and various further examples of pattern selections and languages. These dependencies define relations between patterns, for instance specializations or aggregation. An example for specialization is that a “radio button” defined as pattern is a specialization of the “item selection” pattern defining the selection of items in a UI on a more abstract level.

![Fig. 1. Processing of XML documents using the newly developed extension for MediaWiki.](image)
A Wiki is a tool that is mainly based on presenting one page and thus one pattern at once, which does not facilitate the understanding of the relationships between patterns. Therefore, another extension has been implemented providing a graph-based visualization of pattern dependencies to offer a simplified way to understand these dependencies. The visualization is interactive in the sense that the user is able to change the overall layout of the graph by moving nodes around and to finally get redirected to a pattern’s Wiki page by clicking on the node representing this pattern in the graph. A screenshot of this extension is shown in Fig. 2.

In conclusion, the above-introduced extensions for the MediaWiki engine fulfills the discussed requirements, because they enable the user to query, create, and modify patterns in a Wiki, where the patterns are stored as XML-based documents. The extension uses XSLT styles sheets to prepare patterns to be rendered with the standard Wiki engine based on WikiText and offers extensibility by simply adding further style sheet transforming pattern descriptions into a different target format. This aspect makes the approach usable in various contexts using the process that is presented in the next section.

Fig. 2. Screenshot of the graph view showing the dependencies of the pattern entitled “Extras_on_demand” with other patterns in the Wiki

3 Pattern-based (Semi-)automatic Generation of User Interfaces

The aimed-at process for pattern-based user interface generation can be seen in Fig. 3. The main idea is to use the productive, self-controlling, and collaborative nature of Wikis for the creation of pattern descriptions on the one hand and the feature that combines this repository with transformation approaches to open it up to various approaches for semi-automatic user interface generation on the other, addressing the user groups of researchers and UI developers. The patterns created can be discussed, reviewed, and revised through the collaborative work enabled by the Wiki. The created patterns are based on XML and thus can be transformed using XSLT style sheets to various other XML-based or non XML-based descriptions (such as Java Swing Code), which are finally forwarded to the generation process.

Thus, the process supporting this approach can be divided into two main parts: (a) the collaborative generation of patterns by pattern experts and (b) the use of the pattern’s solution descriptions by a UI generator, such as shortly discussed in the work of
Engel et al. [5] or being further processed to input formats for other generation engines, such as SUPPLE [13].

The result of step (a) could be a pattern as it is shown in the PLML code snippet, below. Here, the pattern “Search Box” as has been defined by Welie.

```xml
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="pattern.xsl"?>
<?xml-stylesheet type="text/xsl" href="XSLT:PLML" title="wikitext" alternate="yes"?>

<pattern xmlns="http://www.hcipatterns.org"
    xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xmlns:html="http://www.w3.org/1999/xhtml"
    patternID="search-box"
    collectionID="welie">
    <name>Search Box</name>
    <problem>The user needs to find specific information.</problem>
    <solution>...</solution>
    <evidence>
        <example>[http://www.google.com www.google.com]</example>
    </evidence>
    <related-patterns>
        <pattern-link patternID="search" collectionID="welie"
            type="" label="Search"/>
    </related-patterns>
    <management>
        <author>Martijn van Welie</author>
    </management>
</pattern>
```

Fig. 3. Reuse process of patterns modeled in the pattern Wiki.

1 www.welie.com/patterns/showPattern.php?patternID=search
The pattern is comprised of a name, a problem, a solution, an evidence, a related pattern, as well as a management node. Most of the XML nodes are self-explanatory. The evidence node specifies examples and rationales showing the pattern’s applicability in existing implementations. The related-patterns node specifies relations to other patterns by pattern-link nodes. These pattern links are then evaluated for the visualization extension. The solution node specifies the solution associated to the pattern. As discussed before, this solution can be defined in different formats. For instance, a solution could be given as XUL snippet as shown below.

```
<solution>
  ==Simple XUL Example==
  <window title="Search Box" orient="horizontal"
xmlns="http://www.mozilla.org/keymaster/gatekeeper/there.is.only.xul">
    <label value="Search for:"/>
    <textbox maxlength="8"/>
    <button label="Search"/>
  </window>
</solution>
```

The XUL solution is labeled by a section headline indicated by == as is common for WikiText of the MediaWiki engine. It is further identified by the namespace specification for the window node that specifies the window being a XUL window. The following part of the solution is given as XUL description of a window with a horizontal layout and is comprised of a label, a textbox for entering the search paraphrase and a button for starting the search. The resulting window can be seen in Fig. 4.

Fig. 4. XUL example for the Search Box example.

In the same way, a ConcurTaskTree could be entered into PLML for the Search Box pattern. This task model can also be extracted by a XSLT style sheet and sent to the discussed PaMGIS engine presented by Engel et al. [5]. Furthermore, XUL or other descriptions can be used to transform those descriptions to other descriptions or code as has been discussed above. Finally, it should be mentioned that a parameterized style sheet could also influence the final presentation of the solution description adapted to a specific use case.

In conclusion, taking the process shown in Fig. 3 into account, the collaborative modeling approach of patterns in combination with a formal and reusable description language opens up a new approach to the use of patterns in the generation of user interfaces. The gap between modeling and use/generation of UIs is reduced and patterns can be enabled to be directly used in the creation of UIs.
4 Conclusion and Future Work

We have presented two extensions for the MediaWiki engine enabling XML-based interaction pattern creation and editing. Using this approach, collaborative modeling of patterns is possible, which is also usable in context of various application scenarios enabled by the use of XSLT style sheets. Adding style sheets to the Wiki, the extension can generate XML-based descriptions of UIs directly from solution descriptions in the pattern. The process introduced above sketches a possible scenario for using this approach supporting the pattern-based semi-automatic generation of UIs. This combination of collaborative creation and editing of patterns reduces the costs for the creation and maintenance of pattern collections and vice versa enhances the accessibility of these kinds of collections. Finally, the use of XML as a basic description format enables the use of these patterns in various application scenarios and tools.

As far as this is ongoing research and development work, various aspect of future work can be identified. First of all, the Wiki has to come to life by a broad contribution of patterns described by experts. Furthermore, the discussed process of semi-automated UI generation should be evaluated and tested in real scenarios. Here, the applicability and effort for using the patterns should be estimated in further research. The creating and editing of patterns should be further supported by the extension through a better generation of appropriate HTML forms. Finally, existing pattern collections, such as the one mentioned in the introduction should be entered into the Wiki.

References

Abstract. Musicians discuss how to represent the music in their performance. In previous work, we have developed a web-based system that enables online discussion by commenting on music score. This paper introduces a new extended system – Co-Musicator that supports musicians to record their performance and make good use of it in their discussion. The system is being implemented as a web application using recent HTML5 technologies. We state various future prospect of the system for better usefulness.

1 Introduction

Most music ensembles, groups of musicians, discuss musical expression to make their performance more impressive. They consider every detail of the piece of music they perform, i.e. the length, the dynamics, the articulation such as accent or staccato, and so on. They also discuss together what they feel and imagine, and how the music is written by the composer. When they meet together, however, they prefer taking time to practice playing together rather than discussing. Therefore, there is a great demand for a system that enables the discussion about musical expressions asynchronously and remotely.

As a previous work, we proposed a method for discussing musical expression on the web-based system[1]. The system enables distant and asynchronous discussion especially for amateur ensembles that cannot spent enough time for practicing. In the system, users can discuss their interpretation by commenting, drawing symbols, linking videos, and depicting “articulation shapes” on music scores.

However, the previous system did not handle recording of their practices. There is a need to record ensemble practices and use the recorded sound data in discussions. By using the recorded data in discussions, the users can discuss and communicate more smoothly and deeply after the practice. They are also helpful for absentees of practice to learn the details. In addition, the group members may perceive their problems by listening their practice recordings.

This paper presents a new extended discussion system that also incorporates the logging of ensemble practices including the recording sound data of practices. The challenges in designing and implementing the system are as follows:
Decreasing the effort to log practices. When players practicing together, they focus on it. Extra effort to logging and recording should be avoided as much as possible. In our system, the players only have to touch the current location in the music score displayed on the tablet PC before playing a instrument.

Visualizing logged practice data and linking to music scores Logs of practices should be easily used in the discussion system. Currently, we assume a log is an entire sound recording of each performance practice. In discussions, the user must be able to refer to the specific part of it and comment on it. In our system, the recorded sound data is linked semi-automatically to a music score utilizing the user’s touch data during practices. Visualization of each sound data is also shown to users so that they can understand which part corresponds to which part in the music score.

Combining logging practices and asynchronous discussions enables more powerful support for musicians. We designed and prototyped a web-based system that enables ensemble players to record practices easily and use them naturally.

2 System Design

2.1 Overview

Co-Musicator is a web-based system for musicians to record their performance, share the sound data with others, and discuss music expressions. Fig. 1
shows the overview of our system. We basically assume two different scenes — co-located practice and spare time. In co-located situations, musicians perform practice and record its sound into the system. On the other hand, after the practice, musicians in their spare time discuss the performance. The music score and the recordings are linked to each other with attached tags.

Because Co-Musicator is a web-based system, various kind of devices with a browser can be used for running Co-Musicator. However, a relatively large display tablet PC is suited for co-located practice. It can display a music score for performing as usual. In addition, the touch display is especially useful for pointing the current location of performance during practices. Thus performing, recording and discussing will be more seamless.

2.2 Two views of Co-Musicator

Fig.2 shows two views of Co-Musicator: the score view (left) and the recorded data view (right). The score view shows music scores, and various information for performance and discussion. In this view, every information is shown on the music score. Users can put annotations such as texts, musical symbols, articulation shapes, and references to recorded data. This view is basically similar to our previous system, therefore, refer to our previous paper[1] for more details.

On the other hand, the recorded data view shows a list of practice logs where each of them corresponds to one meeting of practice. In this view, a log (a sound data) is visualized as a sound wave. Users can select a point in it to playback the sound data from the specified position.

The users can record practices from both of views. The microphone icon at the top-right corner is the button to start recording. During the recoding, the realtime sound input is displayed at the top.

![Fig. 2. Two views of Co-Musicator](image-url)
2.3 Rehearsal Letters and Tags

In general, it is time-consuming to find the point that you want to listen from a lot of long sound recordings. In our system, the users can easily get there by using tags that links sound data to music scores.

Most music scores have “rehearsal letters” and “bar numbers” (Fig. 3) for helping musicians find quickly a certain point on the score in performance practice. Musicians are accustomed to these letters and numbers, so this system makes use of them. Tags in our system are automatically created by the system by using music structure data, but the users can also create and edit them.

![Fig. 3. Reference image of a rehearsal letter and bar number in existing music score (encircled by red rectangle).](image)

2.4 Recording and tagging sound data

Our system provides easy way to record musical performance. The users only have to access the system with a browser and click/tap the ‘Rec’ button. The system shows frequency spectrum of surrounding sound at the top of the view, so the users easily recognize the state of the recording. Another click/tap of the ‘Rec’ button end the recording. The recorded data is shown in the recorded data view immediately.

Currently, we assume that users record one meeting of practice entirely. Therefore, the users basically only have to click/tap the recording button at the beginning and the end of the practice. However, to estimate the correspondence between each time slot in the recorded sound data and the position in the music score, users are encouraged to add tags to sound data during practices. That is, if the user clicks/taps a tag on the music score just before start playing that part, the tag is also attached to the time position in the recorded sound data. By using those tagging information, the system can estimate the correspondence between the sound data and the music score.

To make it easy for users to understand which position in recorded data is performing which position in music score, the system colors each log of practice
Co-Musicator: A Web-Based System for Musicians

(sound wave) using tags information (Fig. 2). The reference color gradation bar is displayed under the music score. The beginning of the music is colored red. The color gradually changes to purple with getting close to the end of the music. Therefore, for example, if a part of sound wave is colored red, that part must be the beginning or around of the recorded data. The system estimates the color of intervals between tags by using the music structure and its tempo. Furthermore, if the volume of a recorded section is smaller than the threshold, it is regarded as not-playing section, and the section is colored gray.

3 Implementation

We are developing a prototype system using the web technologies such as HTML/CSS, JavaScript and PHP. Only a web browser is necessary for using our system, which is advantageous for multi-device and multi-platform for the future.

To record sound, we use Web RTC (Real Time Communication) API and Web Audio API technology. Web RTC API is a new API of HTML5, that can access camera and microphone stream data. Fig. 4 shows the sound data flow in our system.

![Fig. 4. Sound data flow. The sound data from microphone is sent to Web Audio instance. Web Worker encodes it to WAV in background.](image)

While the system captures sound data, the data is sent to instances of AudioContext for display the frequency spectrum. After finishing recording, the data is transcoded to the WAV format, and sent to the source of the HTML audio element as an URL of a binary large object. At this time, Web Workers API is utilized to process the transcoding asynchronously. A sound wave is drawn by calculating maximum volume every unit time from the buffers data.

A tag is simply represented by a div element, and its innerHTML value and an input element is used to edit a tag name seamlessly. All tag name is stored to an array and referred to make a list of linked tags when a tag is clicked.
Commenting function is implemented in similar method with previous work[1]. All of the recorded data, tags, comments and comment regions are stored to the database immediately by using so-called Ajax. Music score is also stored to database in Base-64 format.

4 Related Work

Web-Based Collaboration A lot of web-based collaborative tools and systems have been created for many purpose. Heer et al. developed sense.us for collaborative information visualization with census data of the U.S. on web[2]. Sense.us provides various annotations on the visualized data for discussion and analyzing. Lang and Minker developed a collaborative web-based help-system[3]. It helps users, especially for novice, achieve goals with an avatar that provides information visually and aurally. Goldman et al. developed Collabode that is a web-based IDE for collaborative coding[4]. Collabode has efficiency for various type of coding and they stated the probability of HTML5. Marion and Jomier implemented real-time collaborative visualization with WebGL and WebSocket[5]. These two new technologies allowing faster and easier way to remotely collaborate with 3D datasets.

Our system enables the users to look at music score, record sound, play it and discuss in web browser. We also utilize new technologies : Web RTC and Web Audio. There is extensive collaboration of musician with this system.

Computer-Supported Ensemble As the study of computer-supported musical performance, Bellini et al. researched supporting of real-time music performance[6]. They developed MOODS: a cooperative editor for musical scores that can automatically synchronize written notes on their sheets of music. Sawchuk et al. developed DIP that enables players to participate their practice at a distance[7]. Akoumanakis et al. developed the prototype toolkit for the purpose of distant and asynchronous ensemble practicing[8]. The system DIAMOUSES records performance every part of the music and enable the users to enjoy distant ensemble by using the recorded data. Gurevich developed JamSpace : a real-time collaborative music environment that provides interaction for novice[9].

Our system does not support music performance directly but supports utilization of their performance. We focused on discussions that is a significant part of ensemble practice, and had an eye to indirect supporting musical performance.

5 Future Directions

Our system is still under development, and we have several future directions. First of all, we have to conduct user study of musicians for qualitative evaluation. We are planning usability testing and long-term evaluation in effective conditions.
It may be also possible to make better use of structure data of music scores. In other words, if the system understand every note, rest, and any other musical symbol, rehearsal letters and bar numbers will automatically be recognized as tags. Comments on music score will point out what they mention in great detail. This function may be implemented with the use of a music engraving language such as GNU Lilypond or MusicXML.

In addition, speech recognition function should be implemented in our system. Now our system is able to record all sound that contains performance, voice and noise but not able to distinguish them. In a practice of a music group, musicians of course perform and make sound but they also talk and discuss about their music. In this paper, we have focused on the performance side. If the system can recognize words from voices, there are various probabilities. For instance, the users can use the system without touching devices. The words ‘rec’, ‘stop’ or ‘add tag A here’ will be triggers of operation and their voices of a discussion will be comments immediately. Though it may require high-precision API or library of natural language processing.

If we can use structure data of music score and speech recognition, the system may automatically maps recording sound to specific part of music score. In that case, tagging is no longer required.

6 Conclusion

We have designed a web-based system that supports musicians to record their performance and make good use of it. In previous work, the users can comment on music score for their discussion. Our new system – Co-Musicator provides several functions for simple recording, linking it to music score and discussion using them. It is implemented only by web technology using recent HTML5 APIs.

References

T. Sakamoto, S. Takahashi and J. Tanaka


A Networked Discussion Tool for Shogi

Yoshikazu Tagashira, Hiroyuki Tarumi, Shinji Toshima, Shinji Ohara, Toshiki Kinuhata, and Toshihiro Hayashi
Kagawa University, Takamatsu, Kagawa 761-0396, Japan,
tarumi@acm.org

Abstract. We have defined a discussion model for shogi, a chess-like game in Japan, and developed and evaluated a discussion tool based on the model. In the model and tool design, we have considered the fact that the discussions are on pure logical, huge, complex, and mutually related problems. Using AI programs for advice is also a special feature of the tool. After an evaluation session by university shogi players, the basic tool design has been validated.

Keywords: Shogi, Discussion Support Systems, Board Games

1 Introduction

Design of discussion support tools has been one of the major research topics for CSCW. In the end of 1980’s, gIBIS[1] was a leading research that gave the design of model-based discussion support. One of the most effective applications of model-based discussion is the design of product specifications (e.g., [2]).

Our application domain is shogi, which is a variant of chess-like games, which is very popular in Japan. Among all chess variants, shogi is the most complex one, mainly because it allows reuse of captured pieces. Many players are discussing shogi games and its standard moves. Especially professional and top-level amateur players are everyday doing discussions and create new standards. However, no good computer-based tool supports are provided. For example, the most popular web site for shogi players, “Shogi Club 24” [3], provides only chatting interface for discussions. In case of “81 Dojo” [4], which is an international service, it has graphical interfaces for discussions, but does not support users to model the discussion and save its records.

Such discussions are categorized into two types. One is kansousen, which means a discussion on each game just after the end of the game by the players. Another type is research on standard moves by professional and top-level amateur players. Since both discussions have a common discussion structure, we will not distinguish them, hereafter.

Because discussions on chess-like games are logical, model-based support tools are considered to be useful. We can suppose that existing discussion tools for design specifications can be applied to discussions for games. However, discussions on shogi is different from those on design specifications. Considering the differences, we have characterized the discussion on shogi as follows.
– It is important to record discussions on a database and make them shared, because all discussions are related to each other and they are all for solving one big problem.
– Support for trade-off management is not so important.
– For the discussion model, declaration of “YES or NO” is not required, but it is important to manage comments on positions and variations.
– What we should model is the graph structure of positions and moves, not participants’ opinions.

Moreover, recent AI programs for shogi are becoming stronger. They are almost as strong as professional players. It is useful to incorporate AI programs in the system and let them consider problems and give comments on them.

We have been developing an integrated network-based environment for shogi [5]. The total system is called SAKURA (Shogi Archives and Kansousen Utilities for Research and Advice). Figure 1 shows the configuration of SAKURA. It has a server system and a connected knowledge base. Players use client software to play games and have discussions (kansousen) for their played games. Other users who did not play the game can also join in any discussion using the client software. Incorporated AI programs are not developed by us but we provide interfaces with which external AI programs can connect with SAKURA.

SAKURA Knowledge Base consists of two databases. One is a game record database and another is a position database. The Knowledge Base serves as the shared database for all discussions.

We have already published a short paper on the support of kansousen [6]. In this paper, we extend it to describe the discussion model design, implementation, and evaluation.
2 Discussion Model

We have defined the discussion model based on the experiences of *shogi* players; two of the authors have long experiences. The model consists of three discussion phases: (1) reproduction of a position to discuss, (2) proposal of an alternative move and its succeeding moves (*variation*), and (3) evaluation of the proposed variation and derived positions. Real discussions are modeled as repetitive applications of the three phases.

Figure 2 illustrates an example of game record (the original move, variation ID=0) and its variations in discussion. If one participant wants to discuss position 2 (W84P\(^1\)), the board status at position 2 is reproduced. Then the participant proposes B26P instead of B68S; B26P and its succeeding W34P-B25P are proposed as a new variation: variation ID=1. At position 1 in variation 1, a sub-variation W85P-B77B is additionally proposed: variation ID=2. After these proposals, comments to evaluate the variations will be attached to each derived position.

3 Tool Design

3.1 Protocols

We have defined application protocols corresponding to the discussion model between the server and clients. Besides the protocol-based computer communication, participants have mutual communication by chatting or voice.

**Reproduction of a Position** This protocol is used when one of the players (say, player A) requests to discuss a particular position. The request is sent to another player (player B). If player B accepts it, the focus of discussion moves to the requested position and it is notified to all participants. Figure 3 illustrates the flow of the protocol.

---

\(^1\) The first letter, B or W, means a move by the black and white player, respectively. Two numbers refer to the X and Y values on the board. The last alphabet expresses the kind of piece that moved.
Proposal of a Variation  We have defined a protocol to propose a variation. The proposal will be made when player A moves a piece on the shared board (Figure 4 shown later). The information of move is sent to player B and observers.

The procedure for confirmation of acceptance/rejection is omitted, because it is too irritating and disturbs the discussion. Informal communication by text or voice chatting will be used to avoid conflict of operations. Observers are not allowed to propose variations.

Other Protocols  We have also designed protocols for commenting, and starting/ending discussions.

Implementation  The protocols include pushing communication from the server to participants. We have adopted WebSocket and Ruby technologies for its implementation.

3.2 Database

“Knowledge Base” in Figure 1 is the storage part of SAKURA. It consists of two databases: a game record database and a position database. The position database is the key of our research that cannot be found other shogi databases.

During discussions, participants can save their comments and proposed variations on the game onto the game record database. Comments on positions are saved on the position database. The comments and variations can be retrieved whenever the participants resume their discussions.

The knowledge base has its application interface (API) with http and XML, which provides services for other subsystems.

3.3 Comments from AI

At the early stage of this project, we asked university shogi players how AI programs should join discussions. They did not like active participation of AI because they liked to consider by themselves for their training. Hence, in our design, AI programs show their decisions only when players explicitly request them.
AI programs for *shogi* are designed to use a standardized representation of their decision. They give an evaluation score for a particular position. The score is a positive number when the black player is considered to have advantage and a negative number for the white player. The absolute value of the score shows how much the advantage is.

In Japan, many AI programs have been developed by many researchers. The SAKURA project does not have its own design of AI, but it gives interfaces to existing AI programs. At this time, we have interfaces to Bonanza Feliz[7] AI program. Other AI programs will be supported in future.

When players need help from AI during a discussion, an AI program will show up in the chatting user interface. It gives the best move at the position with an expected evaluation score for the position after the move.

### 3.4 User Interface

SAKURA’s user interface has been designed with the following policies.

1. We have implemented it as a Java Applet application, in order to avoid installing and version-up tasks by the users.
2. The implementation uses multiple windows. It is because the computing environments, especially the size and number of displays, vary depending on the user. Users can layout windows as they like.
3. Besides the shared game board, a personal (non-shared) game board is provided for each user, for their self-study.
4. A shared tree-chart is provided. It gives a graphic representation of a game record and its variations in discussion.
5. For the communication between participants, a chatting window is provided. We have adopted tpChat [8] and customized it for our purpose.

Figures 4 and 5 are samples of implemented screen shots. The shared board (Figure 4) is a WYSIWIS window shared by all participants of the discussion. In the figure, red texts are translations of button titles. This window provides graphical interfaces for commands to start chatting, to ask an AI program for advice, to give comments, and to show the tree chart window.

The personal board (not shown as a figure) has a popular design of user interface for *shogi* applications, which allows users to move pieces for their personal consideration. It is not shared.

The tree chart (Figure 5) is a shared WYSIWIS window, showing the game record (the main horizontal sequence) and proposed variations (branched sequences). When a variation is proposed, a new node is appended and the new status of tree is propagated to other users. When a user clicks a node in the tree, the user can reproduce the position at the node onto the personal board.

SAKURA also provides a chatting window, which is not shown as a figure. A request of reproduction of a position (Figure 3) can be issued by selecting a position on the shared board, clicking a node on the tree chart, or clicking the “Send Request for Position Reproduction” button on the personal board.
A proposal of variation is shared when a piece is moved on the shared board. The proposed variation is instantly appended on the tree chart.

We provide two methods for recording comments for participants. One is to double-click a chatted text; another is to type “#” at the head of a chatting line. A comment is either for a game record or for a position. The user should select one of them.

When participants need an advice from AI, it can be requested by clicking the “Advice” button on the shared board.

4 Evaluation

4.1 Outline of Evaluation

We had evaluation sessions on January 14th and February 1st, 2014, with totally five participants, who were our university shogi players. All of them have experiences of online discussion using a simple chat program on “Shogi Club 24” service. The evaluation session was conducted as follows. We had requested the participants to prepare a game record played between them. The game record was input to our database. We gave tutorials for SAKURA tools to the participants and let them use the tools to discuss the game. During the session, they communicated only by chatting.
4.2 Comparison with the Existing Environments

After the discussion, we requested the participants to give comments comparing with the discussions on “Shogi Club 24.” All of the participants commented that it was easier to have discussions on SAKURA. The reasons were as follows:

– They could move pieces during the discussions.
– They could view the variations with the tree chart.
– They could reproduce positions from the tree chart or the personal board.
– They could take advices from an AI program.

We did not acquire comments comparing with 81dojo.com, because no participants had experiences discussions there. However, all the comments above can be applied to the comparison with 81dojo.com.

4.3 Evaluation of the Commenting Function

No participants used the commenting function. We asked them why they did not and received the following comments:

1. The tree chart provides enough information.
2. Operation for commenting is bothersome.
3. Comments by professional players would be useful, but our comments would not be.

The second comment is on the user interface and the problem should be solved. The third comment is essential and should be discussed here.

We include discussions by professional players in the application domain of SAKURA. The commenting function is mainly designed for such use, because comments by excellent players are useful to many players. In case of amateur players, they would not be positive to give comments that may possibly be incorrect. We will consider to provide personal database to keep record of personal games and comments for amateur players.
4.4 Evaluation of Comments from AI

Participants positively evaluated comments from the AI programs. We will implement interfaces to other AI programs that is compatible with SFEN, which is a standard notation to represent a position.

5 Conclusion

In this paper, we presented a model of networked discussion for shogi and its communication protocols. It suggests that networked discussion systems have to be designed based on specific application requirements. We also gave designs of databases, user interfaces including tree chart and shared boards, and advising functions using AI programs. We showed its implementation and evaluation by shogi players. The evaluation suggested the design was basically validated by the players, but it still needs improvement of user interfaces. We also have future work on how to design the confirmation of requests from other players, and how to deal with observers.

This design will be basically applied to other board games such as chess or other similar games.

Acknowledgments

We appreciate comments from Ms. Madoka Kitao (lady’s professional shogi player) and Mr. Tomohide Kawasaki (81dojo.com).

This research is partially supported by KAKENHI (25540169).

References

A Proposal of Digital Photo Frame
Sharing Photos and Comments

Satoshi Ichimura

School of Computer Science,
Tokyo University of Technology.
Katakura 1401-1, Hachioji, Tokyo, 192-0982, Japan
e-mail: ichimura@stf.teu.ac.jp

Abstract. We developed TwiPhoto, a digital photo frame for mutual
communication based on sharing photos and twitting on twitter. TwiPhoto
can display friends’ newly submitted photos automatically by simply
starting the software. The Twitter comments attached to the displayed
photo are superimposed on the photo. You can take a picture with your
mobile phone and send it to the central picture bulletin board. Then,
your picture will be displayed on your friends’ digital photo frame ap-
pliance. You can see all of the photos that your Twitter friends newly
submitted. Slide are displayed on a TwiPhoto screen as a slide show.
You can also see Twitter comments attached to the photo. The Twitter
comments are superimposed on the photo.

1 Introduction

In recent years, the number of sale of the electronic photo frame appliance called
"a digital photo frame" has been increasing rapidly. It has a function that dis-
plays the photograph taken with the digital camera on the display of an electronic
device which looks like a photo frame. Generally a digital photo frame has a slot
for a memory card. The image files saved in the memory card are shown as a
slide show.

A digital photo frame appliance is often sold as gift goods to be given to
parents, grandparents and relatives. It is now widely popular and useful for
mutual communication among those people [Lindley] [Truong]. For example, the
product of Ceiva has widely acquired popularity as a gift not only in Japanese
market but Western markets [Ceiva].

On the other hand, the network service called Twitter, one of the most fa-
mous microblogging systems, is popular recently [Twitter]. Each twitter user has
his/her own home page called a timeline, and is allowed to send a short text mes-
 sage of 140 characters or fewer to the timeline. Owner’s timeline shows the text
messages sent by not only the owner but also all of the owner’s friends that the
owner follows. Users can follow their friend unidirectionally or bi-directionally.
The communication through Twitter is said that the burden of users is relatively small. On the other hand, when using a conventional SNS in which bi-directional follow is mandatory, users often used to feel a burden for communication, and sometimes give up using it.

We developed TwiPhoto, a digital photo frame for mutual communication based on sharing photos. You can see your friends’ newly submitted photos by simply starting the TwiPhoto software. Your friends’ Twitter comments attached to the displayed photo are superimposed on the photo.

TwiPhoto consists of one picture bulletin board, a web server application that manages digital photos and comments, and a number of digital photo frames which display the photos and comments posted to the central picture bulletin board. Photos can be taken with the mobile phone or the digital camera and sent to the central picture bulletin board. You can post Twitter comments about the photo shown on the digital photo frame display. Posted Twitter comments are sent to the central picture bulletin board and stored in it, and will be downloaded to all your friend’s digital photo frame and shown like a slide show.

This paper mainly describes the background of TwiPhoto development, system design, and evaluation.

2 Backgrounds

A digital photo frame appliance has been used as new communication means among close friends, family or relatives. Typically, conventional products read the graphics file saved in memory card, and display them as a slide show. However, since this type of product displays only the photos stored in the memory card, when time passed, users get bored gradually.

A new digital photo frame product ”ALBO” from SANYO Electric Co.,Ltd has a capability of wireless LAN and 3G networking of Softbank/NTT DoCoMo [ALBO] [PhotoVision]. Its features enables users to update photos stored in the memory card from a remote place. Photos can be sent from a cellular phone or a personal computer.

However, a commercial digital photo frame appliance like ALBO can only transmit a photograph to a photo frame in a single direction. There are no functions for supporting mutual communication among users.

Twitpic is a famous web service implemented using Twitter API [TwitPic]. Twitpic enables users to post photographs from a cellular phone or a personal computer to Twitter. Twitter only allows users to post a text, so that, when a user wants to post a picture, he/she has to use another service like Twitpic. When using Twitpic service, the picture itself is uploaded to Twitpic and the URL of the uploaded picture is appeared in his/her Twitter timeline.

Since, however, Twitpic is a simple picture uploader, it has no function to collect Twitter photos newly submitted by your friends whom you are following on Twitter.
3 TwiPhoto

In view of the above-mentioned background, we developed TwiPhoto, a digital photo frame for mutual communication based on sharing photos.

TwiPhoto consists of a picture bulletin board, a web server application that manages digital photos and comments, and a number of digital photo frames which display the photos and comments posted to the central picture bulletin board. The functions of the system are enumerated below.

1. You can take a picture with your mobile phone and send it to the central picture bulletin board. Then, your picture will be displayed on your friends’ digital photo frame appliance.
2. You can see all of the photos that your Twitter friends newly submitted. Slide are displayed on a TwiPhoto screen as a slide show.
3. You can see Twitter comments attached to the photo as well. The Twitter comments are superimposed on the photo.
4. You can post a comment for the photograph displayed on the picture bulletin board.
5. You can post a comment from your mobile phone for the photograph displayed on the digital photo frame.

The details are described below:

1. Although the basic specification of our picture bulletin board is almost the same as that of other picture bulletin boards, our bulletin board differs from others in terms of using the follow relation of Twitter for user management. When you post a photograph from your mobile phone, you can send it to the dedicated mail address that was designated from the system. You can add a text title to the photo before sending it.

When you post a photograph from the picture bulletin board displayed on your web browser, you log in to the picture bulletin board by using your Twitter account. Since the bulletin board is compatible with the OAuth authentication standard, when you log in to the bulletin board, you can read and write Twitter messages.

A comment can be described in the upper text box in the bulletin board page, and picture to be posted can be selected through the file selection dialog.

The posted contents are automatically transmitted to Twitter (Fig. 1). The posted photo is first stored in the storage of the picture bulletin board, then, a short URL (in http://bit.ly/xxxxxx style) to the stored photo is generated by the system and the URL is posted to Twitter service with the dedicated hash tag "# twiphoto". Note that the short URL and the hash tag will not be shown on digital photo frame appliance and the picture bulletin board.

When you log in the picture bulletin board, you can see not only the photo that you have posted but also all of the photos that your friends on Twitter have newly posted. The system retrieves the follower list of your Twitter account using Twitter API, collects tweets having # twiphoto hash tag, and extracts, from the tweets, the short URLs (in http://bit.ly/xxxxxx style) to the stored photos.

Then, collected photos and a comments are displayed on the top page of the picture bulletin board in order of time (upper part of Fig. 1). The photo is
Fig. 1. Timeline on Twitter

displayed on left-hand side, and one or more Twitter comments attached to the photo are displayed on the right.

(2) By simply starting the digital photo frame, you can see all of the photos that your Twitter friends newly submitted. The photos are displayed on the photo frame screen as a slide show.

When the digital photo frame starts, it access the picture bulletin board. Then, the bulletin board identifies the owner of the photo frame and returns a list of photos to be displayed on the digital photo frame, then the photo frame shows a series of photos as a slide show.

(3) On the upper part of a digital photo frame, a twitter profile icon and a comment of the user who posted the photo are displayed. If Twitter comments are attached to the photo on photo frame, the photo frame superimposes the
comments. The comments are displayed on the photo in the manner of flowing from the right to the left slowly (upper part of Fig. 2).

(4) When you use the picture bulletin board, you can not only see the photo but also add a comment to the photo. By clicking the "reply" button below the photo, a form for writing comment is appeared. You can describe any comment in the form, and send it by clicking the "send" button. The submitted comments are saved into the database of the picture bulletin board.

(5) It is likely that many users of a digital photo frame, typically elderly people, want to use cellular phone instead of personal computer when writing text comment. Actually there are several cellular phone type keyboards, sold in the market, to be utilized to input text into a personal computer [Mevael]. For this reason, we developed the mechanism in which a text comment to the photo displayed on the digital photo frame can be entered from user’s cellular phone.
When you want to compose a comment on shown in the photo frame device, you press the button shown on the photo frame device, then you see QR Code (a kind of 2-dimensional bar code) on the display. QR Code can be read with a wide variety of cellular phone sold in Japan. When you read QR Code shown on the TwiPhoto screen with your cellular phone, the URL to access to the photo stored in the picture bulletin board will appear. Then when you press ok button, you see a form to enter your comment on the photo. The comments entered into the form will be immediately sent to the picture bulletin board and saved into the database.

The picture bulletin board is implemented by PHP and MySQL.

4 Implementations

Fig. 2 shows the implemented digital photo frame with TwiPhoto function. It was developed with Android [Android] SDK, as an application running on Android smart phone (ex. HT-03A, Xperia, Desire) or Android tablet PC (ex. Nexus 7). TwiPhoto software have a capability to access to the picture bulletin board and to slideshow a series of photos and comments downloaded from the bulletin board.

The TwiPhoto software running on the Android is compatible with the OAuth authentication standard, so that you can use the TwiPhoto software with your Twitter account. When you start a photo frame, an initial screen for OAuth authentication will appear, and the latest photographs and comments will be downloaded to the memory card in your photo frame. Then, the photo frame automatically will start showing photos that were uploaded by your friends that you are following in Twitter. Since the time of a last update time is recorded, only new photos and comments that have not been downloaded to the photo frame will be acquired next time.

When one or more comments are attached to a photograph on display, TwiPhoto superimposes the comments on the photograph. It displays so that comments flow from right to left slowly (looks like video sharing site “Nico Nico Douga”) [NicoNico]. If more than one comment is attached to one photo, TwiPhoto delays slightly the timing of showing each comment, and changes the vertical position of each comment, whereas one comment consists of a Twitter profile icon, Twitter account name, and comment itself (Fig. 2).

A longer comment with many characters flows faster than a short comment. The character font of the comment written by your friends whom you follow on Twitter is larger than that of others.

Depending on the number of comments attached to the photo, the time period to show a photo changes. The more comments, the longer the photo remains. Animation effects, such as a horizontal slide, fade-in/out, are applied when a photo changes. You may change the entire speed of movement of comments if necessary.
5 Evaluations

Table 1. Experimental results (5-point scale)

<table>
<thead>
<tr>
<th>Question</th>
<th>Ave.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1 Integration of Twitter and photo frame was good.</td>
<td>3.5</td>
</tr>
<tr>
<td>Q2 Operability of picture bulletin board was good.</td>
<td>3.4</td>
</tr>
<tr>
<td>Q3 Operability of digital photo frame was good.</td>
<td>3.6</td>
</tr>
<tr>
<td>Q4 Flowing comments on a photo was good.</td>
<td>3.3</td>
</tr>
<tr>
<td>Q5 QR code reading function was good.</td>
<td>3.7</td>
</tr>
<tr>
<td>Q6 Operability of QR code reading was good.</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Table 2. Comments from Experimental Subjects

<table>
<thead>
<tr>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
</tr>
<tr>
<td>It was fun to use a smartphone as a photo frame.</td>
</tr>
<tr>
<td>It was fun to see feedbacks from my friends for photos that I had posted.</td>
</tr>
<tr>
<td>It was good to talk about photos that my friends had posted.</td>
</tr>
<tr>
<td>Bad</td>
</tr>
<tr>
<td>Comments superimposed on a bright picture were hard to see.</td>
</tr>
<tr>
<td>Twitter &amp; direct message should have been displayed on a photo frame, too.</td>
</tr>
<tr>
<td>Photos posted to TwitPic should have been displayed on a photo frame, too.</td>
</tr>
</tbody>
</table>

System evaluation was conducted on eight undergraduate/graduate students, who have experience of SNS, in the laboratory of our faculty. All subjects know one another well, and they regularly have a conversation in the laboratory. A Twitter account for every subject was created, and every subject was requested to follow the others on Twitter.

First, subjects were asked to post one or more photos, pictures and comments to the picture bulletin board. Then, each of them was given a digital photo frame, and requested to see some photos on the digital photo frame freely. Then, he/she was asked to post one or more comments for a displayed photo.

To perform QR code reading test, the mobile phones (including iPhone, Xperia, a Softbank terminal, AU terminal, and a Docomo terminal) sold in the market were used. When QR Code is shown on the screen, subjects were asked to read it with their cellular phone, and press ok button, so that they can enter comments on the photo in the photo frame device.

Questionnaires were collected from the subjects after the experiment. Evaluations were done in 5-point scale, where the meaning of each scale was 5 for Yes, 4 for weakly Yes, 3 for neither Yes nor No, 2 for weakly No, 1 for No. The results are shown in Table 1.
Free comments were also collected from each subject, and the collected comments were classified into good opinions and bad opinions. The results are shown in Table 2.

The result in Table 1 describes the average score was between 3 and 4 for all of the questionnaires.

Q1 was to test the usefulness of introducing Twitter communication into a photo frame. Although the average score of eight subjects was 3.4, the average of the two people who regularly use Twitter was 4.0. It is expected that the score would be high if many Twitter users use TwiPhoto.

Q4 was to evaluate the comment posting from a mobile phone. The average score was 3.7. It can be judged that the mechanism of generating QR Code was useful.

Q2 was to test the operability of a picture bulletin board, and Q3 was to evaluate the operability of a photo frame. As Compared a picture bulletin board with a photograph frame, the average score of a photograph frame was slightly higher. The picture bulletin board is considered that there is room for improvement. Although a photo frame got higher score, however, there were two or more bad opinions that required an improvement of the display style of the digital photo frame.

6 Conclusions

In this paper, we proposed a method to integrate SNS and a digital photo frame. This method is designed to encourage communications among close friends, family or relatives. The result of the evaluation suggested our design was useful.

References


Learning Log Navigator: Supporting Authentic Learning Using Ubiquitous Learning Logs

Kousuke Mouri, Hiroaki Ogata, and SongRan Liu

Faculty of Information Science and Electrical Engineering, the university of Tokushima, japan
{mourikousuke, hiroaki.ogata, lb90518}@gmail.com

Abstract. This paper describes self-directed and authentic learning environment supported by "Learning Log Navigator", which is function in SCROLL (System for Capturing and Reusing of Learning Logs). The SCROLL system allows learners to record and share their daily experiences as ULL (ubiquitous learning logs) with locations, photo and/or videos using their mobile devices, and provides quizzes that are generated from ULL. However, learning of answering quizzes is not effective than practical learning in the learning environment. Therefore, this paper describes how to carry out the experience as a task and create experience in authentic learning. In addition, this paper describes how to carry out the experience as a task, create experience in authentic learning and analyze their experiences.

Keywords: Ubiquitous learning logs, learning experience, task-based learning, authentic learning.

1 Introduction

In recent years, the smartphone market is growing rapidly and globally [1]. Currently, Smartphone have been equipped useful tools such as multi-touch interface, full browser, GPS sensor, Acceleration sensor and NFC (Near Field Communication). The owners of a smartphone can be available the learning information and geographical information of around them using their wide various function. Using such function of smartphone, SCROLL (System for Capturing and Reusing of Learning Log) [2] has been developed in order to record and share learning experiences as ULL (ubiquitous learning logs) with locations, photos and/or videos in their daily lives. The system was proposed to record more practical ubiquitous learning logs to improve LORAMS (Linking of RFID and Movie System) [3]. In addition, the system will support international students who learn Japanese in japan using ubiquitous learning logs that they have learned.

For example, international students take a memo what they have learned in their daily lives. However, they do not actively remind themselves of what they have learned, nor the situation where they learned them if the notes have not been taken in detail. Therefore, such issues were improved utilizing the SCROLL system. To date, the system provides some quizzes to recall what they have learned for them. The
quizzes are automatically generated by the system and including such as yes and no quiz, text multiple-choice quiz and image multiple-choice quiz using context data. However, it is not enough to remember what they have learned for a long time. For example, when they learn experiences in their daily lives using SCROLL, the system will record a learning log as snippets of knowledge.

Therefore, this paper proposes how to learn experiences in authentic learning using learning experiences as a learning task. We have been developing learning function called learning log navigator [4] to implement the authentic learning system. Consequently, the system can be aware of learning log around them and navigate them to authentic learning area. However, the system is not enough to recommend or support appropriate learning experiences. Therefore, in order to support or recommend learning task and creating new experiences, this paper describes authentic learning based on learning experiential theory.

2 Related Works

2.1 Task-based learning

Task-based learning has been studied by many researchers in educational engineering. Owing to the advancements in computer and network technologies, the researchers have developed computer-based or web-based learning system for conducting interactive activities. For example, Hwang and Chang or Furtak and Ruiz-Primo have reported the effectiveness of using prompts in guiding students to cope with difficult learning tasks [5],[6]. During the learning activities, individual students are equipped with mobile device for interacting with the learning system. These learning system will provide prompting to learn in the field in accordance with individual students.

We also found some researchers on task-based learning such as Supporting Classroom activities with the BSUL system [7]. BSUL aims to link inside and outside of the classroom by assigning each student with learning tasks. The teacher should create learning tasks in advance and then assign them to the students. When the student carries out the assigned task, he/she can ask the teacher any question online while the teacher can answer students’ questions and assign new task to the student who have finished his/her tasks.

However, task-based learning on these studies is conducted by teacher who created the task. Besides the learning based on these studies can be conducted in short-term. Therefore, in this paper, we propose a learning log navigator that can be assigned learning task in long-term.

2.2 Experiential Learning Theory

Representative learning theory on the task-based learning is a learning process based on experience and it is advocated by Kolb [8]. In addition, incidental and informal learning such as WPL (Workplace Learning) based on Kolb’ experiential learning is advocated by Marsick and Watkins [9]. These theories present a cyclical model of learning, consisting of four stages shown in Figure 1.
1. Concrete experience: The first stage, concrete experience, is where the learners actively experience an activity such as a lab session or field work. It is important that learners go to the learning environment using task.

2. Reflective observation: The second stage, reflective observation, is when the learner consciously reflects back on that experience. During the observation, it is necessary to reflect different viewpoint for learners in order to find new knowledge or correct learning method.

3. Abstract Conceptualization: The third stage, abstract conceptualization, is where the learner attempts to conceptualize a theory or model of what is observed.

4. Active experimentation: The fourth stage, active experimentation, is where the learner is trying to plan how to test a model or theory or plan for a forthcoming experience.

Learning Log Navigator will be provided with learning experiences based on these learning processes shown in Figure 1.

2.3 SCROLL

With the evolution of the mobile device, People prefer to record learning contents using mobile devices instead of taking memos on paper. Most of the language learners have their own learning note as shown in Figure 2. In this paper, learning log is defined as a recorded form of knowledge or learning experiences acquired in our daily lives.

SCROLL has been developed for supporting international students in Japan to learn Japanese language from what they have learned in formal and informal setting. It adopt an approach of sharing user created contents among users and is constructed based on a LORE model which is shown in Figure 3.

In this study, we have developed a learning system called Learning Log Navigator using the SCROLL, to support task-based learning in authentic learning.
3 Learning Log Navigator

We have developed the Learning Log Navigator to improve issues of previous study. There are issues and improvements below on previous studies such as BSUL system and SCROLL.

- Issues
  1. Instructors or administrators must manage and monitor all tasks.
  2. It is difficult for instructors or administrators to provide many tasks for a long term.

- Improvements
  1. Learners will record concrete experiences on the system, and then the system using location data guide them to the authentic learning environment.
  2. When learners reflect on themselves based on reflective observation in the authentic learning environment, the system will provide experiences that others learners have learned.
  3. Supporting to conceptualize from concrete experiences.

To tackle these issues and improvement, Learning Log Navigator aims to automatically provide appropriate learning experiences in accordance with individual learner.

3.1 How to create concrete experiences?

At the beginning, learners need to experience situation, such as shopping in the market, seeing a doctor in the hospital, having a haircut in a barbershop, visiting the museum and so on. However, they might be impossible to immediately grasp geographical information around them.

Therefore, the first stage, they will refer to several learning tasks that other learners have learned. For example, if they would like to learn tasks such as "Let's go to the supermarket to buy tofu", "Let's go to museum to learn art" and so on, they will enter up the learning keyword on the web interface. Consequently, they might be found a
learning task that they would like to learn. They will create experience shown in figure 4 on the web interface after they learned other learners’ task.

When they create new task, they need to enter up task title, target language, place and previous knowledge on web interface. Next step, they will enter up new task script related the task. In this paper, the task script is defined as learner’s concrete experiences, which is recorded activities of each scene with photo and location. Figure 5 shows the task script of recording experiential flow.

The Learning Log Navigator will guide them to authentic learning environment using mobile device and these experiential data recorded by learners.

3.2 Recommendation in the Learning Log Navigator based on Location

To learn at the actual learning environment without instructor or administrator, it is necessary to provide appropriate learning task by system. These issues are solved by previous study already [10]. For example, if some learners have learned the words like "buy", "want" and "tofu" in SCROLL, the system will automatically recommend learners’ related tasks by these words such as the task of "Let’s go to the supermarket to buy tofu". Also, the whole flow of recommendation and analysis of Learning Log Navigator is described as follows.

1. The system will firstly recommend the task near the learner. This is based on the latitude and longitude which are saved in the task by mobile device’s GPS information.
2. The number of task will be controlled by system. If the number of recommended task is too much, learners will be confused. This is because it is difficult for learners to select the correct task. Therefore, the max number of recommended task is controlled by the system until 10.
3. If learners would like to learn more many tasks, not only the system but also the learners can get them by changing the limit of distance. Learners can learn tasks using mobile device through these whole flow.

3.3 How to carry out a task?

At the beginning, Learning Log Navigator will provide learners with appropriate task by analyzing theirs’ learning logs. For example, when learner carries out the task shown in Figure 6, they can get them using android device. Figure 7 shows the task on android device run by Learning Log Navigator, its sample of task is named “Let’s go to the supermarket to buy tofu”. If learners carry out the task, they have to grasp geographical location around them. The interface for solving this problem is shown in Figure 8. After learners select a task, the content of task and geographical information will be verified one time by them. If they believe that they can complete it, they can use to be guided to the learning area. The system will receives the GPS information from learners’ devices and guide the way to destination of supermarket for learners.

Finally, after learners arrive at the destination using mobile device, Learning Log Navigator will be recommended with the learning logs of relating task and the learn-
ing logs around destination. For instance, when learners learn an example tofu through learning activities shown in task scripts, the system will provide some word like Natto, Green soybeans and image that is relevant to the circumstances from SCROLL. Besides, task scripts include communication and conversation in their daily lives. For example, if learners are using a task of “Please ask shop staff the price of tofu”, the task will tell them how to say the sentences by voice, so they can communicate with staff. Also, the knowledge that learners didn’t know will be recorded at the note field on the devices. This is because they share with other learner, like that they can solve the problem with each other.

Based on these interface, Learning Log Navigator will be automatically provided appropriate task.

4 Conclusion and future work

In conclusion, this paper indicated that how to learn using Learning Log Navigator based on experiential theory. In addition, Learning Log Navigator uses the log of international students in SCROLL to share with each other, and guide international students to experience real life Japanese learning activities by the word they learned.

In future work, we will aim finding experiences of effective learning by analyzing many experience acquired in the authentic learning environment. Moreover, we are planning to analyze using algorithm like Tf-idf (term frequency-inverse document frequency), apriori and network centrality [11][12][13]. Also, we believe that finding new pattern of learning or an effective algorithm while analyzing several approaches.

References