

BRINGING THE WIKI COLLABORATION MODEL TO THE TABLETOP WORLD

Stefano Baraldi, Alberto Del Bimbo, Alessandro Valli

Dipartimento di Sistemi e Informatica - Università degli Studi di Firenze
stefano.baraldi@unifi.it, delbimbo@dsi.unifi.it, valli@micc.unifi.it

ABSTRACT

We present an interactive workspace that integrates wiki collaboration in knowledge-building activities with face-to-face scenarios like brainstorming or problem solving sessions. A wiki serves as the repository for knowledge elements, which are presented to co-located users in form of a concept map projected on a table. A computer vision module tracks multiple hands and fingers on the table surface and map elements can be manipulated using a simple gesture language. Also smart devices, like PDAs or tablets, can be connected to the system and participate in the interaction with their local input. The concept map and the wiki are synchronized in real-time, providing notifications to both co-located and distributed users and allowing a community shared awareness that enhances and enriches the knowledge building experience.

1. INTRODUCTION

The natural ability of web communities to be self-organized and self-maintained was at the basis of the *wiki* web invention. A *wiki* [1] is a public repository in form of a web site that can be edited by anyone, where every change to the contents is logged and document versions are archived and available to the users, providing the community a shared awareness on the knowledge base evolution.

These simple characteristics led to a wide proliferation of wiki communities where the *knowledge building* (KB) [2] activity is the focus of user interactions and the hypertext model becomes the natural container of multimedia objects. Wikis are now the core tools of corporate and scientific knowledge management practices, where they become a living memory of the community internal processes, but they are also used in transient collaborative events, like conferences and meetings. In these scenarios a wiki can be used as a “back channel” to promote discussion and create knowledge artifacts that could persist after the event.

In these scenarios the inherent *distributed* nature of wikis is in contrast with the *co-located collaboration* that is going on: users must use a personal device to access the knowledge base and temporarily isolate themselves dividing their attention between the wiki and the live event.

While this is not necessarily a negative outcome, in some collaboration scenarios like briefings, brainstorming or problem solving sessions, active user participation is needed to achieve wanted results. It is therefore interesting to research how the successful distributed collaboration model of the wiki web could

be brought to the mentioned face-to-face knowledge-building scenarios.

The category of CSCW tools that deals with the co-located collaboration is called Single Display Groupware (SDG) [3] and it is designed on the WYSIWIS (What You See Is What I See) paradigm for which users attention should be focused on a single area where session-specific data is presented and simultaneously accessed by means of multiple-user interaction techniques.

There already exist some commercial SDG products [11][12] that integrate the use of multiple remote input devices with middleware components to augment the interactivity of office productivity suites, by adding sketches and notes in top of documents, but they do not deal with knowledge-building activities and feature a classroom-style interactivity. These choices, while useful in some contexts, can be very distant from the natural way of conducting brainstorming and other activities where spontaneity and democracy are the main characteristics and interaction takes place around a table.

Electronic whiteboards like SmartBoard [4] and novel tabletop devices like DiamondTouch [5] provide touch-driven input events, this new modality is leading to the development of a new class of applications based on the wall/tabletop interactivity model.

We think that, in order to bring the wiki collaboration model in a co-located collaborative session, a new kind of workspace and workflow must be designed: featuring tabletop interactivity and a simple and intuitive visual language through which multiple users could spatially arrange and link knowledge elements. At the same time the workspace must keep the visual structure synchronized with on-line repositories accessed by distributed users, following the needs of the KB community.

This paper describes the Natural Knowledge Building (NKM) workspace, a hardware and software infrastructure we implemented in order to reproduce wiki-style collaboration in face-to-face sessions. The interaction metaphor is based on *concept maps* [6], a simple and expressive cognitive formalism, as an approach to the visual management of knowledge structures. Co-located users, physically located in the same room, in front of a wall-sized display or a table, can directly manipulate the knowledge elements (Concepts and Relations) synchronized with a *wiki*. The interaction can happen either using personal devices (PDAs, Tablets) as input methods or through *hand gestures* performed on the presentation surface.

In order to realize a setup that could be easily installed in existing meeting rooms or conference lounges without expensive equipment, the bare hand interaction is accomplished with *computer vision gesture recognition and analysis*.

Vision-based interactive systems are unobtrusive to the users, allowing the realization of a natural interaction, and normal tables or walls can be easily turned into cheap interactive surfaces.

2. THE NATURAL KB WORKSPACE

Wiki collaboration is based on the incremental editing of a collection of web pages, having a meaningful name coincident with the page title, and an hypertextual content providing detail and links to other wiki pages and external resources. While this interface is well suited for distributed users, co-located users typically need a different modality for interacting with data, one that has a higher level of abstraction and is less bound to textual input. This is because in face-to-face sessions the main communication channel is verbal and visuals serve as an aid for symbolic reasoning.

In general, *cognitive maps* are the natural way of handling knowledge visually, and the *concept mapping* formalism showed particularly useful and effective in knowledge building activities [7]. Within our workspace concept maps are used as a visuospatial container of knowledge elements, suitable to portray views of knowledge model subsets.

Every wiki page is focused on a different topic represented by the page name, in this way it can be considered a *concept*, while the hyperlinks to other pages present in the page contents correspond to *relations*. The actual textual content of the page, instead, is not directly mapped and shown in the high level map structure.

Wiki	Concept Map
page name	concept
page hyperlinks	relation
page content	-

Table 1: mapping between wiki and concept maps

Face-to-face session participants use maps to spatially arrange concepts and visually “draw” the relations between them constructing a form of directed graph and effectively enriching the wiki hypertextual model with spatial meta-data and relation labels (section 2.4). Through this *map-building* activity users can also introduce new concepts and relations at a visual level and make it available to distributed users. The wiki users can then visualize the maps and use them as a navigation tool to visit a particular *view* of the knowledge model, detailing and integrating elements introduced in previous sessions.

The multi-user interactivity needed in face-to-face sessions is realized using the SDG component of the workspace called *wikiTable*.

2.1 The wikiTable

The *wikiTable* is a SDG infrastructure, consisting of a hardware and software setup installed in a shared social space that visually presents *Concept Maps* on a flat surface, typically a normal desk or table. The system can manage multiple simultaneous inputs coming from different devices in order to make the map interactive. Session participants can interact in two different modalities: using *remote devices* and manipulating elements with their own *bare hands* on the table.

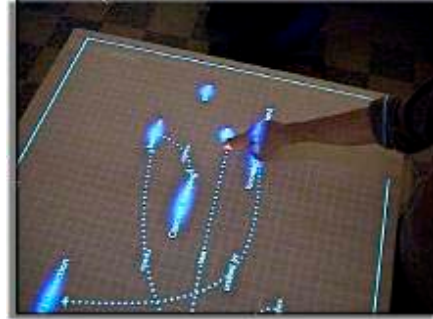


Figure 1: the wikiTable SDG application

Participants can connect to the system using Smart devices (like tablets, PDAs or laptops) equipped with a wireless interface and capable of providing a “pointer” input. Users can join a session through a web page served locally by the embedded *wikiTable* web server and have their personal pointer projected to the interactive surface. In this manner they are free to move around the table while maintaining their attention focused on the single shared map. The local PDA screen is also employed as a mini-interface through which some actions on the map elements can be engaged, visualizing menus that are dependent on the current pointer position and therefore keeping the shared map interface clean of graphical widgets.

Users close to the table can also use *hand gestures* to interact with the map elements, gaining a richer and more expressive manipulation language compared with the “point and click” GUI paradigm and a more *natural interaction* with the visual objects whose affordance becomes similar to that of real objects. Input flows coming from both modalities are continuously merged and the input events are analyzed by device-dependent interaction languages. The outputs of this phase are the following *map-manipulation actions* carried on the visualized elements:

- Add and delete Concepts.
- Draw, modify and delete Relations between Concepts.
- Spatially Arrange Concepts and Relations labels.
- Move the Map viewport

2.2 Bare hand interaction with the Map

In order to manipulate visual objects directly, user’s hands and fingers are recognized and tracked using the *gesture recognition* component of the *wikiTable*, by means of computer vision algorithms applied to the video camera stream of the table area. The input events are interpreted from the posture of multiple hands moving on the table, their orientation and the state of the stretched-out fingers. In this way a simple language has been created using the three recognized postures that are described in table 2.




Posture 1	Posture 2	Posture 3
		
create and modify Relations	grab and move Concepts and Relations	scroll the Map

Table 2: postures used in the map interaction language

Concepts and Relations are represented by their own textual labels. Concepts are created by selecting them from the wiki knowledge repository or by entering text directly, this is accomplished using smart devices text input or a virtual keyboard widget which recognizes key presses of single-fingered hands on key areas. The created words can instantly become Concepts (creating a page in the wiki) or remain in a neutral state, this way they can be used also as spatial notes.

Concepts and Relations can be moved around grabbing their labels with the two-finger gesture. Once grabbed, the object can be controlled only by the user starting the action and its colour changes to indicate the engagement state. A change in the gesture causes the action to end and the object to be positioned in the reached position.

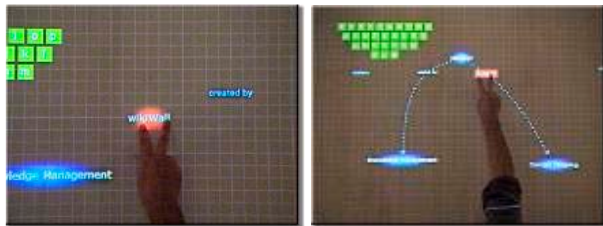


Figure 2: moving a concept (left) and a relation (right)

Relations are created engaging a Concept with a single-finger gesture and moving the created arrow until it collides with another Concept. The resulting line is a spline curve with a single control point in the middle represented by the Relation textual label. Existing relations can be modified in a similar way, grabbing their final arrow and moving it next to another Concept.

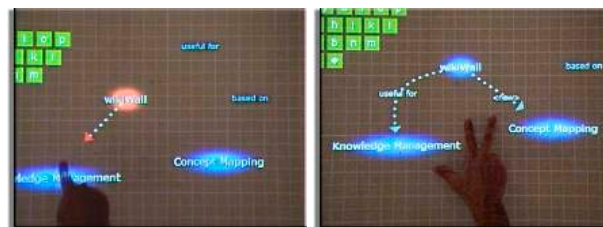


Figure 3: drawing a Relation (left) and scrolling the Map (right)

Relation labels can be assigned grabbing a free word and bringing it on the relation label region. Finally, the Map viewport can be scrolled, allowing the building of large maps, with a specific three-fingers gesture where the map position under the hand is used as a reference point for the scrolling.

2.3 Synchronization with the Wiki

The *wikiTable* embeds a component dedicated to the two-way synchronization with the wiki repository. The protocol is composed of simple directives to query and edit the wiki's status and contents, implemented with a XMLRPC interface over HTTP.

During co-located sessions Concept Maps are created as special wiki pages containing the map structure in XML format. In the wiki web interface these pages are then visualized in vector graphics format through a custom XML to SVG conversion (figure 4). In this way also distributed users can use the maps for

navigating to the pages containing the Concepts details. Relations between Concepts made explicit in maps augment the knowledge model and the hypertextual relations drawn in maps are visualized with icons linking to the related map page. This gives wiki users an extra level of meta-data useful for the knowledge building process.

If page *A* includes a link to page *B*, and one or more maps exist in which relations between *A* and *B* have been defined, links to the involved maps are presented next to the normal hyperlink to *B* and the relation name is shown as a tooltip on the map link. Wiki systems support tags that expand in the page text and layout. Custom tags can be defined to embed in the page panels in which all the relations from/to *A* are listed, introducing another navigation tool.

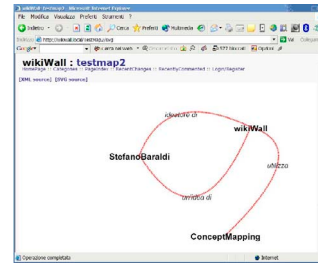


Figure 4: maps are rendered in the wiki through an XML to SVG conversion that retains their visual layout.

Face-to-face users are notified about asynchronous activities performed by distributed users. When the wiki pages are accessed or changed, visual notification are presented on the *wikiTable*. New Concept and Relation elements are shown as visual objects slowly fading away, in this fashion co-located users can consider concept integration in the current map. Other kinds of visual notifications inform session participants that a distributed user has logged in or out of the wiki.

3. GESTURE RECOGNITION

The *wikiTable* setup embeds a computer vision system to perform the needed gesture recognition process. In our setup a ceiling-mounted overhead camera targets the table area and captures video that is analyzed by means of image processing techniques. As in [8] the area is lit by an infrared illuminator and the camera is equipped with a near infrared (830nm) pass filter, in order to capture only the users limbs and not the projected graphical interface, allowing a simpler analysis phase.

The image is first segmented in background and foreground regions through a common *background subtraction* algorithm and the obtained foreground mask is processed by a *finger recognition* algorithm based on [9][10] for a quick early classification of finger point candidates on the whole foreground mask. For the subsequent phases we developed our own *association*, *tracking* and *classification* techniques.

Fingers are tracked over frames with a *nearest-neighbour* association approach extended with position prediction using a simple weighted linear predictor, this technique helps reducing the number of mismatches between different hands' fingers. Hands are then created clustering compatible fingers according to their position and direction, and applying heuristics and estimations on the hand morphology.

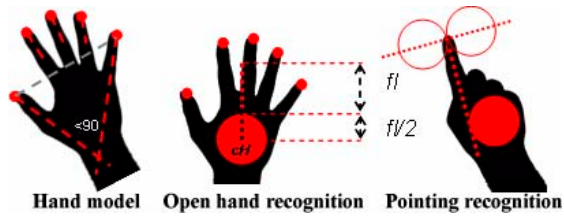


Figure 5: model used in hand recognition

As many as five different postures can be detected for every hand, corresponding to the number of stretched-out fingers, and an arbitrary number of hands can be recognized and tracked over time on the table surface.

Postures are the input atoms of the hand interaction grammar, through which different gestures (the transition of different postures over time along with specific conditions) can be recognized. Through a scripting engine the system allows the creation of many different interaction languages to trigger the application events. We tried to minimize the number and complexity of gestures choosing a simple three-posture language for map interaction, as previously described.

4. RESULTS AND FUTURE WORK

We tested the bare hand interaction system with more than thirty people of different age and computing skills during a six months period. Users were given the task to arrange a reference map from a spatial layout to another. A quick training on the interaction language was provided, in which users were shown the interface usage. Then we observed the video recordings of the map-building session and for every action we determined the average *success ratio* (measuring the times in which the action was successfully recognized by the system) and the average *time duration* (indicating the average time spent to complete the action). For the duration we measured time elapsing from the first moment in which the user intention was clear, up to the time in which the task was successfully completed. These measures are in direct relationship respectively with the effectiveness and efficiency of the system expounded.

We proceeded with two test scenarios: a single user scenario, and a co-located scenario in which three users were asked to complete the task. The results of the observations are shown in table 3.

action	avg. success ratio		avg. time duration (sec)	
	single	multi	single	multi
create label or concept	98%	91%	6.7	8.4
create relation	94%	87%	5.6	7.5
move object	99%	93%	5.2	5.9
scroll viewport	82%	63%	6.3	15.3

Table 3: results of user observations

Single users managed to quickly arrange the map and were only annoyed by the time persistence of 500ms to simulate the click. Multiple users were at first reluctant to interact simultaneously, but then found the co-located map-building experience to be exciting. The only one-user action, viewport scrolling, created some confusion. We believe that a careful design of tabletop system should avoid such exclusive actions.

With the proposed Natural Knowledge Building workspace and the *wikiTable* SDG system, we addressed the integration of two different collaboration contexts: face-to-face sessions and web-based knowledge building using *wikis*.

Further work is ongoing, concerning the definition of other map-manipulation actions (like zooming and map nesting) and a real-time synchronization between remotely placed tables in a scenario where multiple groups of co-located users are geographically displaced when working on the same maps.

5. REFERENCES

- [1] W. Cunningham and B. Leuf, *The Wiki Way*, Addison-Wesley, Boston, MA, USA, 2001.
- [2] Scardamelia, M. and C. Bereiter, "Computer Support for Knowledge-Building Communities." *The Journal of the Learning Sciences*, 3(3): 265-283, 1994.
- [3] Stewart, J., Bederson, B. & Druin, A. "Single display groupware: a model for co-present collaboration", *CHI'99 Conference Proceedings*, 286-293, 2003.
- [4] Smart Technologies. *Smart Board*. <http://www.smattech.com/>
- [5] Dietz, P. and Leigh, D. "DiamondTouch: A Multi-User Touch Technology". In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*. UIST '01. ACM Press, New York, NY, 219-226, 2001.
- [6] Novak, J.D. *Learning, creating and using knowledge: Concept maps as facilitative tools in schools and corporations*. Mahwah, NJ: Lawrence Erlbaum, 1998
- [7] A. J. Cañas, G. Hill, R. Carff, N. Suri, J. Lott, T. Eskridge, G. Gómez, M. Arroyo, R. Carvajal, "CmapTools: A Knowledge Modeling and Sharing Environment", *Concept Maps: Theory, Methodology, Technology, Proceedings of the First International Conference on Concept Mapping*, Universidad Pública de Navarra: Pamplona, Spain. p. 125-133, 2004
- [8] T. Starner, B. Leibe, D. Minnen, T. Westyn, A. Hurst, and J. Weeks. "Computer Vision-Based Gesture Tracking, Object Tracking, and 3D Reconstruction for Augmented Desks", in *Machine Vision and Applications*, Vol. 14, No. 1, pp. 59-71, April 2003.
- [9] J.Letessier, F.Bérard, "Visual tracking of bare fingers for interactive surfaces", in *Proceedings of the 17th annual ACM symposium on User interface software and technology (UIST'04)*, 2004
- [10] C.Hardenberg, F.Bérard, "Bare-Hand Human-Computer Interaction", *ACM Workshop on Perceptive User Interfaces*, 2001
- [11] *Interwrite Meeting Suite*, <http://www.gtcocalcomp.com/>
- [12] *MeetingWorks*, <http://www.entsol.com/>