A2D2A2D… SEAMLESS TRANSFORMATIONS FROM ANALOG TO DIGITAL WORLDS IN SUPPORT OF GLOBAL TEAMWORK

Renate Fruchter

Project Based Learning Laboratory, Stanford University, USA

ABSTRACT: In today's communication intensive environment the rapidly changing nature of work, learning, and play is driven by more and more collaboration, globalization, digital media, interactive devices and spaces, mobility, and convergence of virtual and physical spaces and places. The people and their knowledge are the key corporate asset. Managing, transferring, and reusing knowledge can lead to greater competitive advantage, improved products, and more effective teamwork. The most effective means of knowledge creation and transfer from experts to novices in both education and industry settings is through stories and dialogue using analog channel such as verbal discourse, gestures, annotations, and sketching. Current knowledge capture and reuse solutions do not afford to capture and utilize the relevance embedded in these multimodal streams of communication. This paper explores innovative approaches to support (1) seamless transformations from analog and digital worlds, and (2) cross-media retrieval and interactive replay of multimedia content in support of global teamwork.

KEYWORDS: analog, digital, knowledge management, dialogue, gesture, sketch, data mining, cross-media capture retrieval and replay.

1 INTRODUCTION

What do the Altamira cave paintings, kids' drawings, and professional paper napkin sketches have in common? They all tell a story, but there is no voice of the storyteller. Observations show that the most striking means of knowledge transfer from experts to novices in both education and industry settings is through the informal recounting of experiences from past projects and collaborative dialogue connecting ideas and solutions. Stories convey great amounts of knowledge and information in relatively few words together with sketches on paper or annotations on formal printed or electronic documents. Managing, transferring, and reusing knowledge can lead to greater competitive advantage, improved products, and more effective teamwork. However reuse often fails, since knowledge is not captured, it is captured out of context rendering it not reusable, or there are no formal mechanisms for finding and retrieving reusable knowledge. The digital age holds a great promise to assist in knowledge capture and re-use. The more digital content is created the more paper we print and use. Most digital content management today offers few solutions to capitalize on the core corporate competence, i.e., to capture, share, and reuse business critical knowledge. Digital archives store formal documents (CAD, Word, Excel, etc.) that can be easily edited, shared, searched, and archived. Knowledge reuse and externalization of tacit knowledge is not revealed by these formal documents. The knowledge creation takes place in informal concept generation and problem solving sessions in which knowledge workers gather around diverse documents (e.g., blueprints, contractual agreements, spec sheets, cost estimates, schedules, etc.) and engage in dialogue, annotation, paper and pencil sketching. Paper has a tactile feel; it can be easily folded or rolled and carried to meetings or site visits. It affords single or multiple users to interact and jointly annotate one or multiple documents, and more importantly; it is socially and legally accepted [Sellen 2001]. It provides a high resolution for navigation through the content that enables users to view at a glance local details and global context. However, paper is difficult to modify and expensive to distribute, archive, search, retrieve, and reuse.

In order for knowledge to be captured and reused, the participants should be able to (1) create and express ideas and solutions using natural idioms and channels as communication media such as dialogue, annotation, and paper and pencil sketches, (2) then retrieve, review, and understand the context in which this knowledge was originally created, and (3) interact with the content in a rich, multimedia environment.

Analog communication, i.e., verbal discourse, gestures, paper and pencil sketches and annotations are hard to capture and index in context. Current ICT solutions such as Flickr and YouTube, engage users to provide semantic tags to digital content, i.e., pictures and video. This paper explores innovative approaches developed over the past five years in the PBL Lab at Stanford to support automated and integrated:

- seamless transformations from analog to digital worlds, i.e., analog verbal discourse, gestures, paper and pencil sketches and annotations; digital audio, video, digital sketches and annotated documents, and
- cross-media retrieval and interactive replay of multimedia content in support of global teamwork.

In the PBL Lab at Stanford, we view knowledge capture, sharing, and reuse as key steps in the knowledge life cycle [Fruchter and Demain, 2002, 2005]. Knowledge is created as project stakeholders collaborate on design projects using data, information, past experience and knowledge. It is captured, indexed, and stored in human memory or digital archives. At a later stage, it is retrieved and reused. Finally, as knowledge is reused it is refined and becomes more valuable. In this sense, the archive acts as a knowledge refinery.

Ethnographic studies performed over the past decade of cross-disciplinary teams at work show that a primary source of information behind design decisions is embedded within the verbal conversation among designers. Capturing these conversations is difficult because the information exchange is unstructured and spontaneous. In addition, discourse is often multimodal. It is common to augment speech with sketches as an embodiment of the mental model, or launch into a problem solving discussion triggered by a sketched solution.

This paper addresses the following research questions:

- What are governing interaction and communication principles when a group of project stakeholders engage in a reflective conversation with the situation?
- How can such multimodal communicative events expand a building information model (BIM) to become a rich multimedia building knowledge model (BKM)?
- How can the analog and digital worlds be bridged to support a seamless transformation from analog to digital to analog to digital (A2D2A2D…) in support of the knowledge life cycle?

2 POINTS OF DEPARTURE

The topic of how to capture knowledge in project teams has received extensive attention from researchers in design theory and methodology. The value of contextual design knowledge (process, evolution, rationale) has been recognized. Nevertheless, the additional overhead required of the user in order to capture it has precluded these applications to be widely adopted. Other studies focused on either the sketch activity, i.e., learning from sketched accounts of design [Tversky 1999, Guimbretiere, 2003, Stiedel and Henderson 1983, Olszweski 1981, Kosslyn 1981, Goel 1995] and verbal accounts of design [Cross 1996, Cross 1992, Dorst 1996]. Researchers studied the relation between sketching and talking [Eastman 1969, Goldschmidt 1991]. Recent studies of interactive workspaces [Ju et.al, 2004] explore capture and navigation issues related to technology-augmented interactions. To help guide the users' exploration of an archive of unstructured dialogue and sketch content linked to structured, document databases, it is necessary to develop a search and retrieval mechanism. The reported research builds on Donald Schön's concept of the reflective practitioner paradigm of design [Schön 1983]. Schön argues that every design task is unique, and that the basic problem for designers is to determine how to approach such a single unique task. Schön places this tackling of unique tasks at the center of design practice, a notion he terms *knowing in action* (Schön 1983, p. 50). To Schön, design, like tightrope walking, is an *action-oriented* activity. However, when knowing-in-action breaks down, the designer consciously transitions to acts of reflection. Schön calls this *reflection in action.* In a cycle which Schön refers to as a *reflective conversation with the situation*, designers reflect by *naming* the relevant factors, *framing* the problem in a certain way, making *moves* toward a solution and *evaluating* those moves. Schön argues that, whereas action-oriented knowledge is often tacit and difficult to express or convey, what *can* be captured is *reflection in action*.

The following working definitions for *data, information,* and *knowledge* are used in order to formalize governing principles of team communication. Data (e.g., printed documents or digital documents of CAD, spreadsheets, text) represent the "raw material." This is easy to manage and store in corporate databases or ftp sites. Nevertheless, data is not information. Information emerges during a communicative transaction between a sender and a receiver. Information is created as the sender takes data and adds meaning, relevance, purpose, value through a process of contextualization and synthesis. Neither data nor information represents knowledge. Observations show that knowledge is created through dialogue during one's thought process or among people as they use their past experiences and knowledge in a specific context to create alternative solutions. During these dialogues knowledge is created as connections, comparisons, combinations, and their consequences are explored. It is important to note that documents do not reveal the tacit knowledge externalized during such communicative events. The documents also ignore the highly contextual and interlinked modes of communication in which people generate concepts through verbal discourse and sketching. This exploration views the act of reflection in action as a step in the knowledge creation and capture phase of what we call the "knowledge life cycle" [Fruchter and Demian, 2002, 2005]. Knowledge represents an instance of what Nonaka's knowledge creation cycle calls "socialization, and externalization of tacit knowledge." [Nonaka and Takeuchi 1995]. We build on these constructs of the knowledge lifecycle and the "socialization, externalization, combination, and internalization" cycle of knowledge transfer.

Current research studies present media specific analysis solutions, e.g., Video Traces for video content annotation [Stevens, Cherry, and Fournier, 2002], Tracker for video content processing based on object segmentation (SRI) [SRI, 2002], Fast-Talk for audio search [product of Fast-Talk company], text vector analysis and latent semantic indexing for information retrieval from text repositories [Landauer and Dumais, 1995] [Salton, Buckley, and Singhal, 1995], video object segmentation of video footage [Farin, Haenselmann, Kopf, Kühne, and Effelsberg 2003], [Kuehne, 2002]. Nevertheless, there are few studies focusing on cross-media capture and retrieval.

3 SEAMLESS TRANSFORMATIONS FROM ANALOG TO DIGITAL WORLDS

Ethnographic observations of collocated and geographically distributed project teams were used to identify and model the activities, interaction, and information generated by the stakeholders. A scenario-based approach [Rosson and Carroll 2001] was used to design new ICT mediated human-to-human interactions. The premise behind the scenario-based method is that descriptions of people's interactions are essential in analyzing how technology can support and improve their activities. The objective was to identify governing principles of analog communications and map them into a digital interactive environment.

Observations in the analog paper world show that during communicative events there is a continuum between discourse and sketching as ideas are explored and shared. A primary source of knowledge behind decisions is embedded within the verbal conversation among stakeholders. The link between dialogue and sketch provides a rich context to express and exchange knowledge. This link becomes critical in the process of knowledge sharing, retrieval and reuse to support the user's understanding of the shared information and assessment of the relevance of the retrieved content with respect to the task at hand. Nevertheless, paper is a media hard to share, exchanged, and re-use, and does not capture the discourse among users. The moment you lost the paper sketch the ideas are lost.

The scenario based analysis of collocated and geographically distributed project teams led to:

1. the formalization of a *problem space* defined by three dimensions:
   - number of participants – from single to multiple participants,
   - number of artifacts (paper or digital) – from one to multiple artifacts (e.g., documents, models, etc.) and
   - number of input devices used to sketch or mark up digital or paper documents and manipulate models – from one to multiple input devices (e.g., pens, markers).

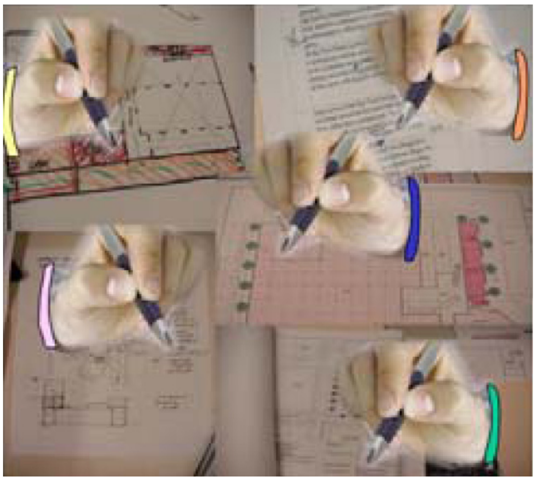
   This problem space defines a spectrum of interaction scenarios of increasing complexity. These interaction scenarios are consistent with the observed communicative events in real project teams and work settings. For instance, the two extremes of the spectrum are defined by the following two interaction scenarios:
   - at one end of the spectrum: a single participant interacting with one document or artifact, using one input device or pen, and
   - at the other end of the spectrum: multiple participants interacting with diverse documents or artifacts using multiple input devices.
2. the definition of a *reflection in interaction* model. The reflection in interaction model focuses on multiple participants engaged with a project situation that extends Schön's concept of reflection in action that describes a single participant interacting with the situation. The interaction scenario defined by "a single participant marking up one document with one input device/pen" matches Donald Schön's concept of *reflection in action* of a single practitioner [Schön 1983]. In the case of the *reflection in action* a single practitioner has a reflective conversation with the design situation. This entails the following activities:
   1. Naming the relevant factors in the studied design,
   2. Framing the problem in a specific domain,
   3. Making moves towards a solution, i.e., often modifying the design solution to address some of the identified problems, and
   4. Evaluating the moves or proposed modifications.

Each move or modification made by one team member in one discipline can impact solutions in other disciplines (e.g., a change made by the architect in the floor plan layout can impact the structural system solution proposed by the structural engineer). This in turn creates a new situation for that team member and triggers a reflection in action cycle in that domain.

The second interaction scenario defined by "multiple participants marking up multiple documents with multiple devices/pens" matches the concept of *reflection in interaction*. As participants review concurrently multiple documents, they have a constant *reflective conversation with the situation, the artifacts or documents,* and the other *participants*. Their interactive reflective process consists of the following activities:

1. Identifying the relevant factors in all considered disciplines through exploratory sketching and discussion.
2. Correlating these factors across disciplines and documents.
3. Discussing and exploring alternatives across disciplines.
4. Assessing alternatives and their implications.

(a)



(b)

Figure 1. Interaction Scenarios: (a) *Reflection in Action* – defined by a single participant marking up one document page with one input device/pen, and (b) *Reflection in Interaction* - defined by multiple participants marking up and correlating multiple documents with multiple devices/pens.

It is interesting to note that whereas action-oriented knowledge is tacit and difficult to transfer, what can be captured and transferred is the action itself in relation to the reflection in interaction that reveals the rationale and correlation across disciplines and documents, as well as the new knowledge that is created through discourse among the stakeholders. Capturing, sharing, and reusing knowledge created in cross-disciplinary, collaborative teams is critical to increase the quality of the product, reduce hidden work (i.e., less coordination and rework [Levitt and Kunz, 2002]), improve communication and knowledge transfer among the stakeholders, decrease response time and decision delays, reduce time-to-market, and cost. The objective is to reduce the number of requests for information cycles to one cycle.

Two prototypes that complement each other have been developed, tested and deployed to achieve the goal of knowledge capture, sharing, and reuse through seamless transformations from analog to digital and digital to analog worlds: TalkingPaper[1] [Fruchter et al 2007] and RECALL[2]. [Fruchter and Yen, 2000]. Both prototypes model the concepts of reflection in action and reflection in interaction.

TalkingPaper is a client-server environment that bridges the paper and digital worlds. It leverages technologies such as Anoto^{TM} paper, digital pens (e.g., by Nokia, Logitech, Maxell), cell phones, Bluetooth communication between digital pen and cell phone, and GSM/GPRS network services. Talking paper system (Figure 2):

- provides an analog-to-digital content conversion that enables seamless transformation of the informal analog content, such as dialogue and paper and pencil sketches, into digital sketch objects indexed and synchronized with the streamed digital audio of a TalkingPaper session. This conversion takes place in real-time, with high-fidelity, and least overhead to the participants. Sketches and annotations on the Anoto^{TM} paper are converted into digital sketch objects that are synchronized with the speech from the digital audio channel and the documents form the corporate database that were printed on Anoto^{TM} pages.
- supports knowledge sharing allowing the user to understand the content in the context it originated, i.e., streamed, interactive replay of indexed digital audio-sketch rich multimedia content that captures the creative human activities of concept generation through dialogue and paper and pencil sketching. The TalkingPaper sessions are automatically uploaded from a the TalkingPaper client to a TalkingPaper Web server that was developed to archive, share, and stream these sessions on-demand. It automatically synchronizes digital audio-sketch episodes with the corresponding document that was printed on the Anoto^{TM} paper used in that session.

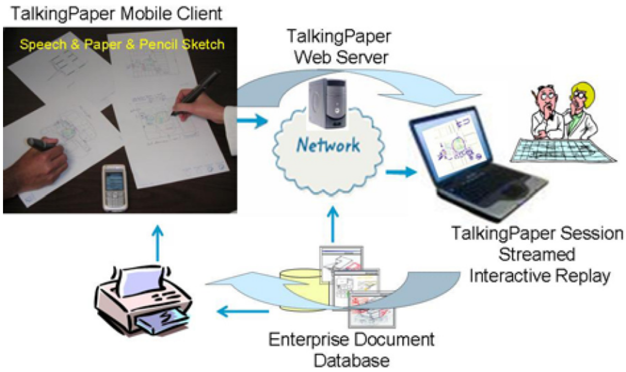


Figure 2. TalkingPaper Prototype bridges Paper and Digital Worlds.

It is important to note that TalkingPaper allows for an endless cycle of transformations from analog paper to digital multimedia content and back to analog paper as TalkingPaper Web pages can be printed for further annotation and discussion, and subsequent posting on the TalkingPaper Web server etc. until a decision is taken by the team members. (Figure 2) TalkingPaper can be used in both reflection in action and reflection in interaction scenarios.

RECALL is a client-server environment that bridges the analog and digital worlds. It leverages technologies such as SmartBoard and Tablet PC for direct manipulation of digital project content and sketching. (Figure 3) RECALL drawing application written in Java captures and indexes each individual action on the drawing surface. The drawing application synchronizes in real time with audio/video capture and encoding. Once the session is complete, the drawing and video information is automatically indexed and published on a RECALL Web server that streams the sessions. It supports distributed and synchronized playback of the sketch and audio/video from anywhere at anytime. In addition, the user is able to navigate through the session by selecting individual drawing elements as an index into the audio/video and jump to the part of interest. Users can create free hand sketches or import any images (e.g., CAD) and annotate them during their discourse. At the end of the session the participants exit and RECALL automatically indexes the sketch, verbal discourse and video. This session can be posted on the RECALL server

1 Provisional patented, PBL Lab, Stanford University.
2 Patent, PBL Lab, Stanford University.

for future interactive replay, sharing with geographically distributed team members, or knowledge re-use in future projects. RECALL aims to improve the performance and cost of knowledge capture, sharing and interactive replay. RECALL has been used extensively in various scenarios, such as, team brainstorming individual brainstorming, best practice knowledge capture.

Both TalkingPaper and RECALL sessions can be accessed with their URL and streamed on IE browser. More importantly, these sessions can be linked to specific graphic objects in a 3D building information model (BIM). This extends the 3D model from a BIM that provides data and information about the future facility to a building knowledge model (BKM) that integrates and shares stakeholders' knowledge created over during the project.

4 CROSS-MEDIA SEARCH AND RETRIEVAL

Once the dialogue, speech, and sketches are captured in the form of indexed and synchronized *digital video-audio-sketch* content by either TalkingPaper or RECALL, the digital multimedia sessions can be shared and replayed by team members anywhere anytime. However, as more and more digital content is archived it becomes difficult and time consuming to find and retrieve relevant footage from such large, unstructured, digital data stores. To address this challenge the PBL Lab team developed the DiVAS (Digital Video Audio Sketch) system. It presents a cross-media semantic analysis and data mining methodology of indexed digital video-audio-sketch content that captures the creative human activities of concept generation and problem solving. DiVAS provides a macro-micro index to large digital archives of rich, multimedia, and unstructured content. Knowledge re-use is facilitated through contextual exploration and understanding of the multimedia content that is retrieved.

The key activities and steps for digital content processing identified in support of effective knowledge retrieval and reuse are - capture, retrieve, and understand. The *capture* phase involves collecting the analog data (dialogue, speech, and sketches), digitizing, indexing and synchronizing it. This step is performed by either TalkingPaper or RECALL. In the *retrieve* step this digital data (video, speech and sketch) is then processed in order to add structure to the unstructured multimedia data (i.e. video, audio and sketch) by identifying key information occurring in the data archive and automatic mark up of digital footage for future search and retrieval. Finally, the *understand* module examines all the structured information gathered from the different streams and creates the context between them, i.e. finding relevance between the information in each of the streams. The DiVAS system integrates the following modules to achieve the cross-media capture, retrieval, and replay to understand knowledge in context: (1) TalkingPaper or RECALL prototypes for capture, (2) I-Gesture prototype for semantic video analysis, and (3) i-Dialogue prototype for text and data mining from the speech channel. The integrated result is a refined and highly contextual cross-media representation of the knowledge captured that is relevant to the specific project query posted in DiVAS by stakeholders.

DiVAS builds on the following observations of communicative events and hypotheses:

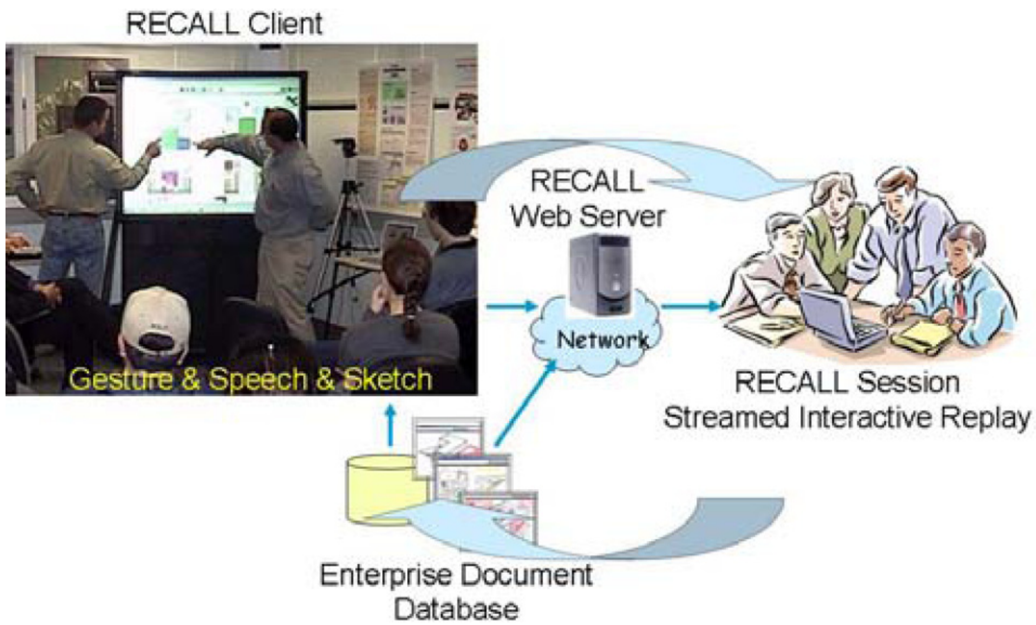


Figure 3. RECALL Prototype

- Gestures can serve as external representations of abstract concepts which may be otherwise difficult to illustrate. Gestures often accompany verbal statement as an embodiment of mental models that augment the communication of ideas, concepts or envisioned shapes of products. A gesture is also an indicator of the subject and context of the issue under discussion. If gestures can be identified and formalized they can be used as a knowledge indexing and retrieval tool, they can prove to be useful access point into unstructured digital video data. I-Gesture methodology and a prototype allows users to (1) define a vocabulary of gestures for a specific domain, (2) build a digital library of the gesture vocabulary, and (3) mark up entire video streams based on the predefined vocabulary for future search and retrieval of digital content from the archive. [Fruchter and Biswas, 2007] I-Gesture prototype takes advantage of advanced techniques for video object segmentation and automatic extraction of semantics out of digital video. These techniques are used to develop well defined, finite gesture vocabulary that describes a specific professional gesture language to be applied to the video analysis. *I*-Gesture can be used both as an integral part to DiVAS system as well as an independent video processing module
- Speech is a fundamental means of human communication. Design and construction are social activities. Observations and evidence show that designers and builders generate and develop concepts through dialogue. The objective is to mine the captured rich, contextual, communicative events for further knowledge reuse. i-Dialogue methodology and prototype: (1) adds structure to the unstructured digital knowledge corpus captured through TalkingPaper or RECALL, and (2) processes the corpus using an innovative notion disambiguation algorithm in support of knowledge retrieval. [Yin and Fruchter, 2006]. i-Dialogue takes advantage of advanced techniques for voice-to-text conversion (e.g., Naturally Speaking, MS Speech Recognition) and text search. As other studies have shown, text is most promising source for information retrieval. The information search applied to the audio/text portion of the indexed digital video-audio-sketch footage results in relevant discourse-text-samples linked to the corresponding video-gestures.

The DiVAS system provides an innovative cross-media search, retrieval and replay facility to capitalize on multimedia content stored in large corporate repositories. The user can search for a keyword (a spoken phrase or gesture). The system searches through the entire repository and displays all the relevant hits. On selecting a session, DiVAS will replay the selected session from the point where the keyword was spoken or performed. It is important to note that video and audio/text provide a macro index and the sketch provides a micro index to large, unstructured repositories of rich multimedia content. These streams can be also run independently. The background processing and synchronization is performed by an applet that uses multithreading to manage the different streams. We developed a synchronization algorithm that allows us to use as many parallel streams as possible. Therefore there is the possibility of adding more streams or modes of input and output for a richer experience for the user.

The value of DiVAS will be most perceptible when the user has a large library of multimedia sessions and wants to retrieve and reuse only the segments that are of interest to him/her. Currently, solutions for this requirement tend to concentrate only on one stream of information. The advantage in DiVAS is literally 3 fold because the system allows the user to measure the relevance of the query using 3 streams, gesture, verbal discourse, and sketch. In that sense, it provides the user with a true 'multi-sensory' experience. For example, if the user is interested in learning about the dimensions of the cantilever floor, his/her search query (e.g. 'cantilever') would be applied to both the processed gesture and audio indices for each of the sessions. They would serve as a 'macro index' to the items in the archive. Some segments will be identified by an (1) I-Gesture hit that is cross-indexed with speech, text and sketch (shown as a thumbnail image), (2) an i-Dialogue hit marked-up in the text and cross-indexed with video/audio and sketch, or (3) both an I-Gesture and i-Dialogue hit cross-indexed with the corresponding sketch. If there are a large number of hits for a particular session and the hits are from both audio and video, the possible relevance to the user is much higher. In this case, the corresponding gesture could be one related to width or height, and the corresponding phrase could be e.g. 'cantilever' e.g., within a single session or in different sessions. (Figure 4) The cross-media display of the results to the query allows the user to make a more informed decision which is most relevant and replay it to understand and possibly reuse data, information or knowledge from that segment.

To evaluate the performance of DiVAS a series of experiments were conducted. The testbed was the archive of projects from Architecture/Engineering/Construction (AEC) Computer Integrated Global Teamwork course [Fruchter 1999] [Fruchter 2004]. This academic testbed simulates the collaborative design process and the industry environment. Design knowledge has been archived and processed by i-Dialogue and I-Gesture. The archive includes team sessions and individual team member sessions that are consistent with typical design-construction industry setting. The archive consists of imprecise digital content together with closely related precise digital content, such as project websites, project documentations, and messages posted on the project discussion forum. For the precise digital content, more than 500 megabytes of project documents were filtered and mined. Based on the meta data, such as the owner of the document, the time stamp of the document, precise digital documents are associated with the imprecise digital documents as discussed in the chapter of i-Dialogue, so that the precise digital documents can be used in the notion disambiguation process. For the imprecise digital content that includes both voice recordings and video recordings marked up with gesture labels, 30 sessions from the project archive are analyzed and processed for the evaluation experiment. The variables in the experiments are: 1) type of speech transcripts, clean or dirty; 2) whether i-Dialogue is applied; and 3) whether I-Gesture is applied. There are three essential aspects for the evaluation methodology: 1) a benchmark document collection; 2) a benchmark suite of queries; 3) a binary assessment - relevant or irrelevant - for each query-doc pair. The evaluation uses the benchmark document collection of the speech transcripts document collection obtained from the testbed. The benchmark suite of queries is manually generated by the authors who then read through all the perfect speech transcripts as well as relevant project documentation for each session. The keywords or phrases that summarize each session were identified. The queries are constructed from these keywords and phrases. The binary assessment - relevant or irrelevant - for each query-doc pair is constructed by querying the keywords or phrases with the perfect speech transcripts. The evaluation question is: *Will information retrieval over the dirty text have the results similar to the information retrieval over the clean text if the dirty text is updated by I-Dialogue and I-Gesture?*

In order to validate the information retrieval improvement of the cross media relevance ranking, a series of experiments were performed. The speech transcripts archive was modified into six different cases:

- Case #0: "clean" speech transcripts only – manually transcribed speech sessions that are used as the comparison base
- Case #1: "dirty" speech transcripts only – automatically transcribed speech sessions with transcription errors (e.g., "Cantilever" is transcribed as "*can we deliver*.")
- Case #2: "dirty" speech transcripts and notion labels
- Case #3: "dirty" speech transcripts and gesture labels
- Case #4: "dirty" speech transcripts, notion labels, and gesture labels (gesture labels are integrated after I-Dialogue™ is applied)
- Case #5: "dirty" speech transcripts, notion labels, and gesture labels (gesture labels are integrated before I-Dialogue™ is applied)

The query terms are: "architectural constraints", "height limitation", "cantilever", "tension ring", "concrete steel", "bracing", "foundation", and "load path". The query results for these six cases are illustrated Figure 5. The query results are measured using precision and recall metrics. The relevance ranking sequence corresponding to each queried notion for all six cases is compared. There are two ratios used for further analysis: 1) *Recall*, which is the fraction of relevant docs that are retrieved; 2) *Precision*, which is the fraction of retrieved docs that are relevant. The *Recall vs. Precision* curve is drawn for evaluation illustration. The case #0 is used as the comparison baseline. It is located at the top right corner. The case #1 represents the worst case, in which we run query with dirty text only. The query results score about 50% for both recall and precision. Using i-Dialogue only, both recall and precision can be improved to almost 60% - case #2. If only I-Gesture is used, the ratio is improved close to 60% - case #3. However, the precision ratio might be lower compared to case #4, since I-Gesture tends to bring in both relevant and irrelevant documents. If both I-Gesture and I-Dialogue are combined to update the dirty text, the results further improve the recall and precision to 80% - case #4 and #5.

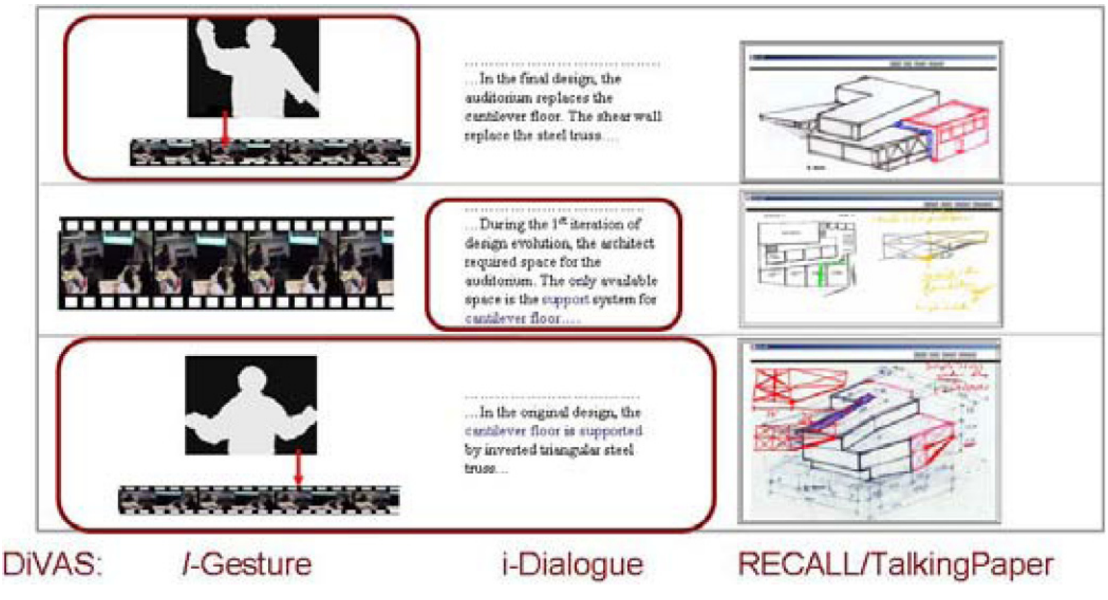


Figure 4. DiVAS Prototype for Cross Media Capture, Search, and Retrieval.

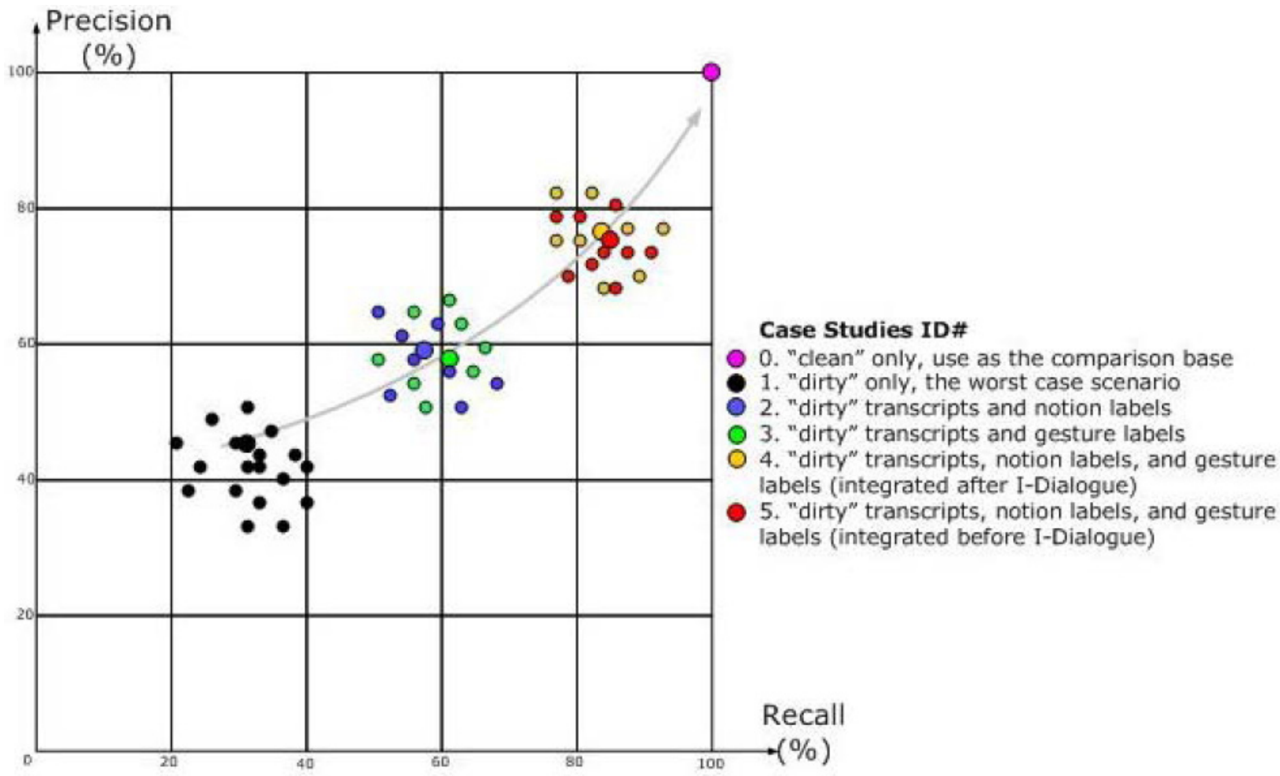


Figure 5. Recall vs. Precision Evaluation Results.

5 CONCLUSION

This paper explores the interaction among participants in a group of project stakeholders and defines governing communication principles. These translate into (1) a three dimensional problem and solution space, (2) a model for reflection in interaction that extends Schon's concept of reflection in action, and (3) requirements and innovative approaches to support (i) seamless transformations from analog and digital worlds modeled and tested by two prototypes – TalkingPaper and RECALL, and (ii) cross-media retrieval and interactive replay of multimedia content in support of global teamwork modeled and tested with the DiVAS prototype. The value of these concepts and prototypes indicates opportunities to (1) capitalize on corporate core competence that resides in its people and knowledge, (2) extend a building information model (BIM) to become a rich multimedia building knowledge model (BKM), and (3) mine the unstructured, rich, multimedia archives that will grow as digital technology becomes a ubiquitous part of work, learning, and play.

ACKNOWLEDGEMENT

The studies and prototypes presented in this paper were partially sponsored by PBL Lab and MediaX at Stanford, Nokia, Cisco, and KDDI. Dr. Fruchter would like to thank her research students who engaged in the R&D projects to develop the ICT prototypes – Zhen Yin, Praktik Biswas, Shubashri Swaminathan, Manjunath Borajah, Chhavi Upadhyay, and Mudit Garg.

REFERENCES

Cross N., Christiaans, H., Dorst K., (1996). *Analyzing Design Activity*, John Wiley UK.

Cross, N. and Roozenburg, N., (1992).*"Modelling the Design Process in Engineering and Architecture*", Journal of Engineering Design v3 n4, pp. 325-337.

Dorst, K., Christianns, H., Cross, N. (eds), (1996).Analyzing Design Activity, Wiley.

Eastman, C.M., (1969) *"Cognitive processes and ill-defined problems: A case study from design*", Proc. Int. Joint Conference on Artficial Intelligence, pp. 675-699.

Farin, D., Haenselmann, T., Kopf, S., Kühne, G., Effelsberg, W. (2003) "*Segmentation and Classification of Moving Video Objects*." Borko Furht, Oge Marques Handbook of Video Databases, CRC Press.

FAST-TALK http://www.fast-talk.com/

Fruchter, R., Swaminathan, S., Boraiah, M., and Upadhyay, C., (2007) "Reflection in Interaction," AI & Society International Journal

Fruchter, R., and Biswas, P., (2007) "Using Gestures to Convey Internal Mental Models and Index Multimedia Content," AI & Society International Journal

Fruchter, R. (2004) *"Global Teamwork: Cross-disciplinary, Collaborative, Geographically Distributed e-Learning Environment,"* Collaborative Design and Learning: Competence Building for Innovation, ed. J. Bento, J. Duarte, M. Heitor, W. Mitchell, Quorum Books Greenwood Pub. Group, Inc., 2004, 265-297.

Fruchter, R. and Demian, P. 2002, *CoMem: Designing an Interaction Experience for Reuse of Rich Contextual Information from a Corporate Memory*, AIEDAM Journal, 16, 127-147.

Fruchter R. and Demian, P. (2005) *"Corporate Memory*," in "Knowledge Management in Construction," ed. by C.J. Anumba; C. Egbu; P. Carrillo, Blackwell Publishers.

Fruchter, R. and Yen, S., (2000)*"RECALL in Action,"* Proc. of ASCE ICCCBE-VIII Conference, ed. R. Fruchter, K. Roddis, F. Pena-Mora, Stanford, CA.1012-1021.

Fruchter, R. 1999, *Architecture/Engineering/Construction Teamwork: A Collaborative Design and Learning Space, ASCE* Journal of Computing in Civil Engineering, October 1999, Vol 13 No.4, 261-270.

Goel, V., (1995) Sketches of Thought, MIT Press, (1995).

Goldschmidt, G., (1991) "The dialectics of sketching", Creativity Research Journal, v4.n2, pp 123-143.

Guimbretiere, F., (2003) *"Paper Augmented Digital Documents,"* HCI Lab, University of Maryland.

Ju, W. Ionescu, A., W. Lawrence Neeley, Brydon, R. Stone, M., and Winograd, T. (2004) "Beyond Meeting Capture: WorkspaceNavigator," CHI 2004

Kosslyn, S. (1981) *"The medium and the message in mental imagery: A theory",* Psychological Review, 88, pp 46-66.

Kuehne, G. (2002) "*Motion-based Segmentation andClassification of Video Objects*", PhD Thesis, Univ. of Mannheim.

Landauer, T. K., & Dumais, S. T. (1995) "*A Solution to Plato's Problem: The Latent Semantic Analysis Theory of Acquisition, Induction and Representation of Knowledge"* Psychological Review, 104, 211-240.

Levitt, R. and Kunz, J. (2002) "Design Your Project Organization as Engineers Design Bridges," CIFE Working Paper WP073.

Nonaka and Takeuchi, (1995) *The Knowledge-Creating Company*, Oxford University Press.

Olszweski, E. J., (1981) *The Draughtsman's Eye: Late Renaissance Schools and Styles,* Cleveland Museum of Art/Indiana University.

Rosson, M. B. and Carroll, J. M. (2001). *Usability Engineering: Scenario-Based Development of Human Computer Interaction*

Salton, G., Allan, J., Buckley, C., Singhal, A. (1995) "*Automatic analysis, theme generation, and summarization of machine-readable texts*", In Readings in Information Visualization, pp. 413-418. Morgan Kaufmann, Los Altos, California.

Sellen, A.J. and Harper, R.H.R., (2001) *"The Myth of the Paperless Office,"* MIT Press.

Schön, D. A., (1983), *The Reflective Practitioner*

SRI Meeting Chronicler http://www.esd.sri.com/projects/mtg_chronicler/

Stevens, R., Cherry, G., and Fournier, J. (2002) *"Video Traces : Rich Media Annotations for Teaching and Learning"* CSCW2002.

Stiedel and Henderson, (1983) *The Graphic Languages of Engineering*.

Tversky, B., (1999) *"What Does Drawing Reveal About Thinking?"* Proc. Visual and Spatial Reasoning in Design.

Yin, Z. and Fruchter, R. (2007) "i-Dialogue: Information Extraction from Informal Discourse," AI & Society International Journal.