# Virtual Time Machine

**Michitaka Hirose, Katsuhiko Sakaue, Ryoko Ueoka**
Research Center for Advanced Science and Technology, the University of Tokyo
4-6-1 Komaba Meguro-Ku, Tokyo 153-8904, Japan
National Institute of Advanced Industrial Science and Technology (AIST)
1-1-1 Umezono Tsukuba-Shi Ibaragi-Ken, 305-8563, Japan
*hirorse@cyber.rcast.u-tokyo.ac.jp,sakaue@aist,yogurt@cyber.rcast.u-tokyo.ac.jp*

## Abstract

This paper describes the ongoing project called Virtual Time Machine. The VTM project aims to realize to experience future and past events by long-term storage of information as well as prediction of future events by simulations using the storage of information.

**Key words**: Virtual Time Machine, Behavior mining, Simulation of possible world, Real-time human simulation

## 1. Introduction

Typical functions of computers are (1) long-term storage of information and (2) prediction of future events by simulations. If these two functions are fully realized, we should be able to experience past and future events as we do present events. To make a bold statement, a computer becomes a virtual time machine.

## 2. Virtual Time Machine Project

A technology-development project based on the above concept, in which we overcome the so-called "time wall" using information technology, is the virtual time machine project (VTM). This project was initiated as a policy-budget project of the National Institute of Advanced Industrial Science and Technology (AIST), and it is a collaborative research effort carried out by AIST with the participation of a group from the University of Tokyo. The first author planned this project as a research coordinator in the field of information communication.

In recent years, as the keyword "ubiquitous" implies, many ideas have been proposed for using computers at any time, in any place, and by anyone. Some of the concepts proposed are the Project Oxygen (MIT)[(1)], Pervasive computing (IBM) [(2)], Cooltown (HP) [(3)], Project aura (CMU) [(4)], Invisible computing (Washington University) [(5)], Location aware computing (Intel) [(6)] and EasyLiving (Microsoft) [(7)]. Furthermore, the idea of storing data on all the activities of an individual using a storage device, the size of the memory space of which is increasing to an almost limitless level, has been introduced. Microsoft's MyLifeBits[8] adopts this idea. The Defense Advanced Research Projects Agency (DARPA)[9] has also adopted this idea. However, no study has been conducted on the development of an information supporting technology which eliminates the barrier of "time" and in which future events are predicted on the basis of past and present events through a combination of sensing, data mining in a real environment and simulation. This project covers the following three technologies.

### (1) Behavior mining

Behavior mining is a technology to comprehensively observe and analyze human behavior. To use a large amount of accumulated data instantaneously, this technology structuralizes behavioral records on the basis of the meaning of each behavior, such as by segmentation. This technology is a compound technology consisting of sensing, archiving and data mining. It is being developed by AIST, and by applying it, dynamic information from people and from society may be observed and accumulated using various sensors for images, voices and tags. It has been improved in terms of accuracy, function and stability to a level at which the technology can always be operable in a real environment. At the same time, data mining and statistical analysis of real-time data obtained are carried out to convert the real-time data into structured information as a resource which enables the secondary understanding, abstracting and retrieving of information, which leads to the ability to predict future events. In concrete terms, the following examples are considered: 1) a monitoring system based on a new concept, which enables the prevention of crimes and accidents by automatically analyzing human behavior using computers; at the same time, people can benefit because they feel that they are protected by the system, and 2) an archive system/analysis system based on human activity records at offices and public institutions. To realize behavior mining, not only a compact and high-speed sensing system that realizes real-time responses, but also an environment for computing and communication must be established, in which an overwhelming amount of data continuously output from sensors for humans and the society is processed.

**(2) Simulation of possible world**

The simulation of a possible world is a technology required to present various possibilities created from human behavioral selection in terms of scenarios. This technology belongs to a compound field consisting of simulation and data mining. The following three items: 1) the generation of scenarios which may be realized in the future, 2) the presentation of different past and present realities which might exist if some conditions had been different (Historical IF), and 3) the possible world simulation technology which supports decision-making and learning in people and society through the verification of the past and present, are realized using the three subsystems listed below.

a) Multiagent simulation:
By integrating a model composed of a group of agents which have functions such as recognition, judgment and learning, and a conventional model of the physical field, the overall possible world is comprehensively simulated.

b) Generation of scenarios:
By integrating the extraction of system structure from real data using data mining and retrieval of past actual case data, past cases which resemble the present state are retrieved and corrected; a future scenario is generated on the basis of the outcome.

c) Ubiquitous agent:
In the ubiquitous information environment, where communication and sensing technologies are integrated, this sub-technology designs and installs (1) a group of agents which offer services in a distributed manner, and (2) an overall architecture. This sub-technology includes mobile phones, car navigation systems, appliances and wireless LAN equipment as terminals and its own infrastructure.

**(3) Real-time human interface**

Real-time human interfacing is a technology to provide a sense of time, such as compression, expansion and sharing of the time axis, which people have not experienced previously. This technology is in a field related to the so-called next-generation human interface, such as the interface of the five senses and cognition interface.

We will develop a new human interface, which can control compression, extension, transfer, retrogradation, creation and sharing of the time axis freely.

In concrete terms, oldness/history and access histories are visualized; visualization technology for declining information is adjusted to the time axis. Research is also carried out on information retrieval systems using (1) the reversible operation of events by zooming of the time axis and (2) neighborhood relationships (such as those between generation time, content, attribute and place). A flexible retrieval method using time information is then developed, and the method is applied to the real world.

In this project, we plan to integrate the three technological items described above into three prototypes.

(1) VTM office

We aim to construct a room/space where time can be rewound. The theme here is to summarize a meeting of approximately 2 hours into 10 minutes and to prepare real-time generation of minutes.

(2) VTM wearable

Support for present and future behavior, as well as support for archiving behavior on an individual level, is carried out. An interface to significantly improve individual behavior along the time axis is developed.

(3) VTM public

We aim to record and predict events in shops and public spaces. Through analysis and prediction using data obtained, various supports for space management, such as disaster prevention, traffic control and marketing may be carried out.

Methods not only to pursue convenience and comfort, but also to secure relief and safety in society using IT technology are targeted as the main fields of application. We prepare to further develop information technology that contributes to society by generating a new technological stream in which a new sense of value based on information technology is introduced against the important time axis that most significantly restricts people and society, while using the prototype development procedure as a trigger. The areas of practical application assumed are follows.

- Contribution to relief and safety: detection of suspicious action/ prevention of crimes, generation of real-time disaster prevention scenarios (evacuation, guidance), systems for prevention of accidents on roads and the prevention of accidents in facilities and in the field
- New functions used in offices: real-time advanced summarizing of minutes, management of meetings using structural summaries, concept support system, advanced judgment support system using the Historical IF
- Optimization in public spaces: control of crowds in theme parks by analyzing group behavior, support for group users, application to marketing and control of the flow of people
- Advancement at operating fields: predictive process control, micromarketing, support for learning in business scenes, support for operating sequences, prediction of dangers for workers
- Personal information support: realization of a predictive personal scheduler which supports instantaneous decision making accommodating individual conditions

Figure 1. VTM Office Image(Meeting)

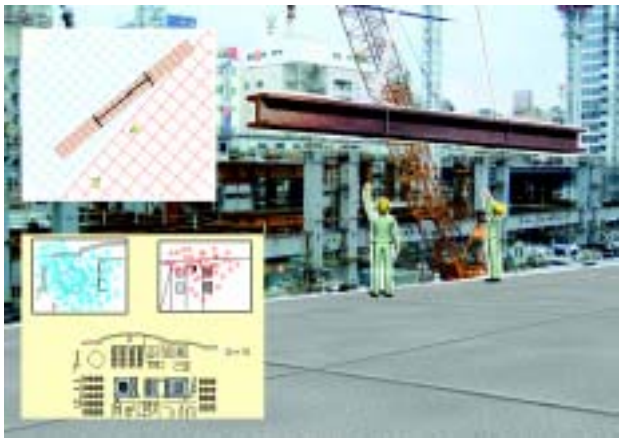

Figure 2. VTM Wearable Image ( Construction site)

Figure 3. VTM public image ( VTM vehicle case)

## 3. Three prototypes

**(1) VTM office**

Research on the generation of minutes using speech recognition has been carried out at several research institutions. However, the word error rate (WER) of such systems is generally approximately 30-40%. The products cannot be directly used without correction or editing by a person. If a meeting is a formal one such as a conference, such assistance by people is meaningful; however, if the meeting is a small-scale one, it is not cost-effective. In addition, the atmosphere of the meeting is not conveyed by minutes.

Therefore, a minute generation system as a digital archive, but not minutes recorded by characters (in print), is considered assuming that the system is used to support meetings. The content of a meeting is recorded and subjected to a speech recognition system; however, the output thus obtained is not directly used as minutes. The result of speech recognition is considered a tag for the recorded speeches and images; it is designed to effectively capture the outline of the meeting. The goal is to understand a meeting of approximately 2 hours in approximately 10 minutes. The entire meeting is not recognized and the data are not handled in the original signal format; instead, signal processing using a format intermediate between the two, i.e., an intermediate coding scheme, is carried out, and the appearance vectors are analyzed initially by statistical processing.

Using a microphone array as an input device, it is possible to put directional tags on the input speech. Using these tags, the minutes are structuralized and visualized. In addition, the changing of speakers (?) can be visualized leading to the effective reproduction of the minutes. In addition, using the sound source separation function of the microphone array, when the dialogues of two people are superimposed, the statements can be separated, stored, reproduced and recognized as two different sentences. Furthermore, by the individual tracking function using images, high robustness against the movement of individuals is realized.

First, the system carries out pretreatments, such as detection of the start of a speech, estimation of direction, separation of sound source, removal of noise and accommodation to noise for each speaker using microphone array processing. Next, rough analysis by dictation using an intermediate coding scheme, a general-purpose language model and key word extraction are carried out. After that, on the basis of the information obtained, the topic is estimated. A language model specific to each topic is selected; then, a second speech recognition and a script for the entire meeting are generated. On the basis of this script, information is retrieved and highlighted. The VTM office presents past speech (past) and predictable conclusions, and other possible conclusions (future). A system carrying out these tasks is to be studied and developed.

Figure 4. Multi-directional camera & microphone array mockup ( VTM office)

**(2) VTM wearable**

We aim to realize a wearable VTM, which not only recalls past memories and conditions by constantly acquiring and accumulating daily histories of movements and activities, but also selects information and services which will be required for an individual based on the prediction of the individual's activity and effectively presents and provides such data. For example, an individual's hobbies and tastes are estimated on the basis of the behavioral history obtained from various sensors, and only the necessary information is automatically selected and presented from the excess information provided by various media such as the Internet. Another application is the use of behavioral histories for health care and to suggest recommended exercise and meal patterns. Furthermore, assuming many people will use wearable VTMs, it may be possible to improve the ability not only of an individual, but also of a group as a whole, through the prediction of the timing and priority of instructions required for field workers.

To realize such wearable VTMs, wearable sensor modules, which can constantly acquire and accumulate data in daily life and operations, are indispensable. In this project, we are developing a small sensor module for monitoring daily movements such as walking and climbing up and down stairs, which are required to grasp the history of the movement and the physical condition of an individual. Furthermore, in addition to conventional HMDs and displays of mobile phones, we are developing an interface to effectively present information provided by the wearable VTM through a tactile display which monitors vibration and a wearable active camera laser (WACL, an expandable real world interface to present information directly to the real world).

By fusing data mining and simulation technologies, the activity level of multiagents related to individual internal conditions is evaluated using an individual's behavior log. On the basis of these data, individual conditions are automatically evaluated, the transition between conditions is estimated, and the habits of an individual are structuralized; in this way, the applicability of a wearable VTM in fields such as health care is demonstrated. In addition, we demonstrate the feasibility of a possible-world-simulation wearable support function, by combining highly efficient retrieval through information filing, capturing conditions, and future estimation.

**(3) VTM public**

Assuming an individual as an agent, experiments on real-environment multiagent simulations are carried out by extracting moving lines and value-added information from an identical field. Because data on moving lines alone are not sufficient as information for data mining and simulation, we will develop a detection method, which enables the detection not only of moving lines but also of the directions of the body and face simultaneously using a stereo camera (ubiquitous television (USV)).

On the basis of the results of behavior mining about the behavior of an individual, multiagent simulation may be carried out to analyze the behavior of a group. Changes in density, moving lines of a group, average residence times, target achievement rate and satisfaction level of an individual are analyzed by changing the placement of various pieces of equipment in the environment and providing information to each individual. Various possible worlds which are related to marketing and controlling of the flow of people may be predicted through simulations while the conditions are interactively changed. At the same time, we tackle the extraction of suspicious behavior by the following method; normal behavior is statistically modeled, and the rest is considered suspicious and the presentation of such movements is examined.

Regarding multiagent simulation technology, we simulate the flow of people and the amount of traffic, propose several algorithms related to providing information and controlling the flow of people, and verify the results in simulations. We constructed consorts as information architecture for linking on-time data in the real environment and a multiagent simulator, to provide information services for users. In this project, assuming the connection of the VTM public to a stereo sensor, research and development in the following two areas are carried out: (1) rule calibration used in simulations by accessing on-time sensor data (preprocessing of calibration) and (2) the interface that interactively changes simulation conditions, verifies the user scenario, and presents results after mining so that the results correspond to the user's scenario (postprocessing of simulations).

## 4. Space and time

In the 20th century, we developed technology related to space to an ultimate level. The development of transportation technology led to the production of aircraft by which the space for our activities has been expanded to most places on earth. Similarly, the development of the technology of architecture has led to the production of high-rise buildings and enabled the three-dimensional use of urban space.

The word "time" is used as a pair with "space." When we consider "time," we recognize that we have not been able to manipulate it at all. As has always been the case, one day consists of 24 hours, which cannot be expanded or exchanged. If we may exaggerate, overcoming time will become a main theme of technology in the 21st century.

## 5. Closing remarks

Today, many people have an absurd impression of time machines. However, the activities of people who work mostly with e-mail are automatically compiled. In

addition, a large number of people optimize their behavioral scenarios using station- and car-navigation systems. Our present daily life would look like science fiction to people 30-40 years ago.

## References


1. http://oxygen.lcs.mit.edu/
2. http://www306.ibm.com/software/pervasive/index.shtml
3. http://www.cooltown.com/cooltown/
4. http://www-2.cs.cmu.edu/~aura/
5. http://portolano.cs.washington.edu/
6. http://www.intel.com/labs/wireless/lac/
7. http://research.microsoft.com/easyliving/
8. http://research.microsoft.com/barc/mediapresence/MyLifeBits.aspx
9. http://www.darpa.mil/ipto/Programs/lifelog/