KANSEI The Technology of Emotion

AIMI International Workshop

Proceedings

Antonio Camurri
Editor

"Interactive Poem"

Naoko Tosa and Ryohei Nakatsu

ATR Media Integration & Communications Research Laboratories

{tosa,nakatsu}@mic.atr.co.jp

Abstract

We propose a new type of speech-based interaction system called "Interactive Poem." Conventional speech-based interaction systems have only focused on the transmission of logical meaning involved in speech. The application of such systems has been restricted to business services such as making reservations or data retrieval. In the Interactive Poem system, however, a human and a computer agent create a poetic world by exchanging poetic phrases, thus realizing Kansei-based communications between computers and humans. This paper first proposes the concept of "Interactive Poem". It then describes the details of the system we have developed, including the software and the hardware configurations as well as the interaction mechanism.

1. Introduction

In human oral communications, Kansei information, such as emotions and sensitivities, plays a very important role. Kansei information is sometimes more important than the logical information included in speech. This can be confirmed by the fact that babies start to recognize emotional information before they can recognize logical information in their mothers' voice. In the case of adults, we too can recognize what other people want to say at a deeper level by integrating both logical and Kansei information included in speech. This is the key to smooth communications. Unfortunately, in the field of AI so far, the focus has been on recognition of only meaning information, while Kansei information has been neglected as noise.

Based on the above considerations, we started to study how to realize Kansei-based communications between computer agents and humans. As a first step toward this aim,

we have developed several computer agents such as "Neuro Baby"[1] and "MIC and MUSE"[2]. These are computer characters that are capable of recognizing several emotions included in speech and reacting to them by changing their facial expressions and body motions. Fortunately, these agents have been very successful and have been demonstrated at various exhibitions.

Emotion-based communications, however, is only part of Kansei-based communications, and we are interested in yet a deeper level of communications. As a next step toward the realization of Kansei-based communications between computer agents and humans, we have selected "poem" as a means of communications. There are several reasons for this approach. The main reason is that in a poem not only the meaning of words or phrases but also the rhythms and moods created by their sequence plays an essential role. Therefore, the poem is intended to transmit Kansei information such as mood and sensitivity rather than logical information. The second reason is that poems were originally expressed by oral reading rather than in writing. This means that a poem is suitable for interaction between computers and humans. Recently, researchets have shown increased interest in the realization of Kansei-based interactions and communications between computers and humans[3][4][5]. However, only few have treated voice communications, despite the fact that voice is an essential means of Kansei-communications. This is the third reason why we are interested in treating communications based on an uttered poem.

This paper first introduces the concept of interactive poem. Then the basic principles of the Interactive Poem system we have developed based on the concept are described. The software configuration and hardware configuration will be

described in detail. Finally, a typical installation of the Interactive Poem system will be introduced.

2. Concept of Interactive Poem

"Interactive poem" is a new type of poem that is created by a participant and a computer agent collaborating in a poetic world full of inspiration, emotion and sensitivity.

In a conventional poem, a poet tries to convey emotions and sensitivity through a sequence of carefully selected words and short phrases or sentences. Because of the magic power unleashed by the words and phrases, people can easily understand the message that a poet wants to express and can thus enter this world created by the poet. However, the world of each poem is static and, therefore, limited by the intentions of the poet because the phrases, sentences and especially their sequence are fixed.

The concept of this interactive poem is based on conventional poetry, but goes beyond traditional limits by introducing the capability of interaction. A participant and a computer agent create a dialogue by exchanging short poetic phrases, and through this exchange produce a new poetic world that integrates the poetic world of the agent with his/her own.

A computer agent called "MUSE", who has been carefully designed with a face suitable for expressing the emotions of a poetic world, appears on the screen. She will utter a short poetic phrase to the participant. Hearing it allows him/her to enter the world of the poem and, at the same time, feel an impulse to respond by uttering one of the optional phrases or by creating his/her own poetic phrase. Exchanging poetic phrases through this interactive process allows the participant and MUSE to become collaborative poets who generate a new poem and a new poetic world.

3. Software Configuration

The system used to create the interactive poem consists of four main units: system control, speech recognition, computer graphics generation and speech output (Fig. 1).
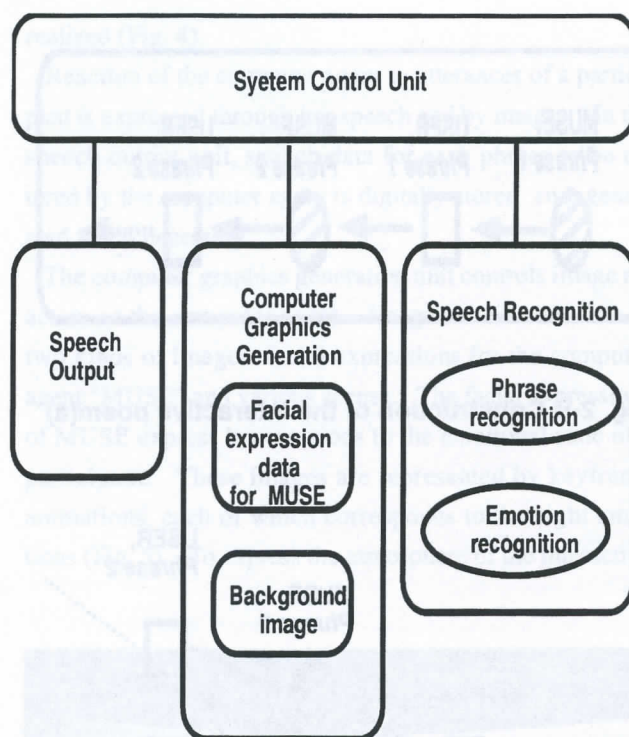
The system control unit manages behavior of the whole

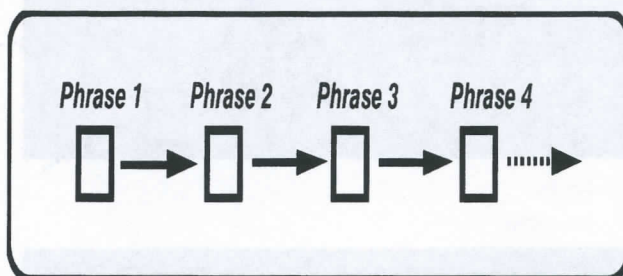**Fig. 1 Block diagram of the Interactive Poem System**

Fig. 2-A Conventional poem

system by utilizing the interactive poem database. In this system, the most important issue is constructing the interactive poem, so we must first explain how the interactive poem database is constructed. A conventional poem is considered a sequence of poetic phrases. In other words, the basic construction of a conventional poem can be expressed by a simple state-transition network where each phrase corresponds to a given state, and for each state there is only one successive state (Fig. 2-A).
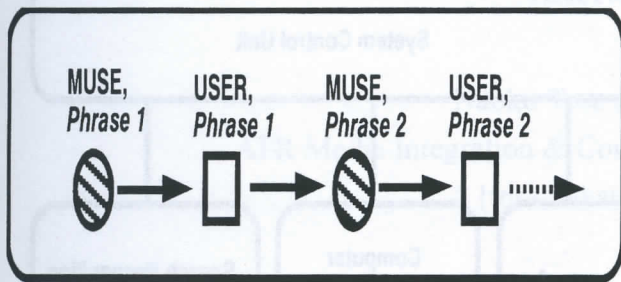
55

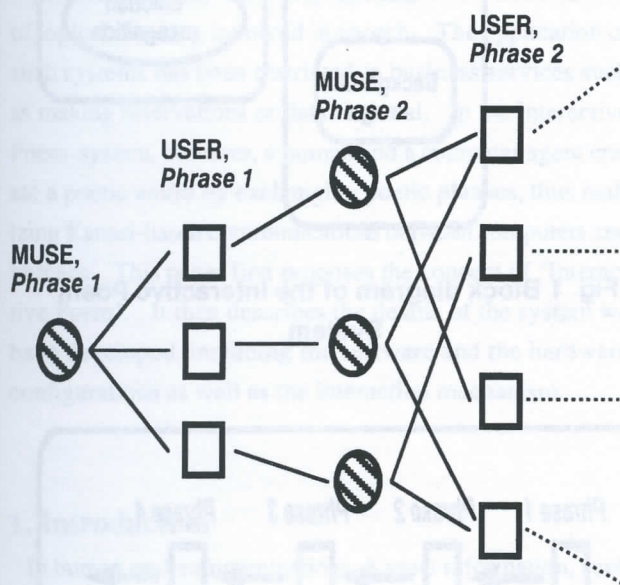**Fig. 2-B Construction of the Interactive poem(a)**

Fig. 2-C Construction of the Interactive Poem(b)

The basic form of the interactive poem is expressed by this simple transition network, but it differs from a conventional poem in that phrases uttered by the computer agent and phrases uttered by a participant appear in turn. This corresponds to a simple interaction where the computer agent and a participant alternately read a predetermined sequence of poetic phrases (Fig. 2-B).

To introduce improvisational interaction into our system, we modified this simple transition so that multiple phrases are connected to each phrase of the computer agent (Fig. 2-

C). These phrases are carefully created and chosen by taking into account how well their rhythms are formed and the meaning of each phrase. This transition network is stored in the interactive poem database and used to control the whole process.

The speech recognition unit has two different speech recognition functions: phrase recognition and emotion recognition. To recognize each phrase uttered by a participant, we have adopted HMM (hidden Markov model) based speaker-independent speech recognition technology. Each phrase to be uttered is represented in the form of a phoneme sequence and is stored in the lexicon (Fig. 3). To simultaneously detect the emotional state of a participant, an emotion recognition function is introduced. A neural network architecture has been adopted as the basic architecture for emotion recognition. This neural network is trained with the utterances of many speakers to express the eight emotional states of joy, happiness, anger, fear, teasing, disgust, disappointment and emotionless. As such, speaker-independent and content-independent emotion recognition is
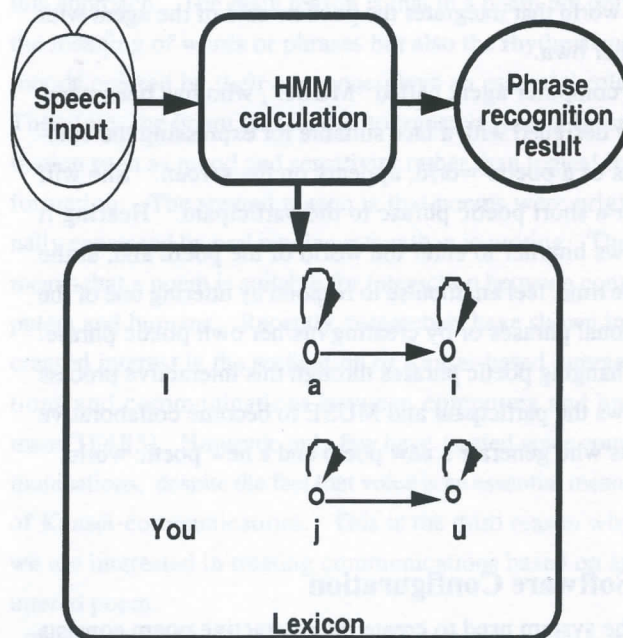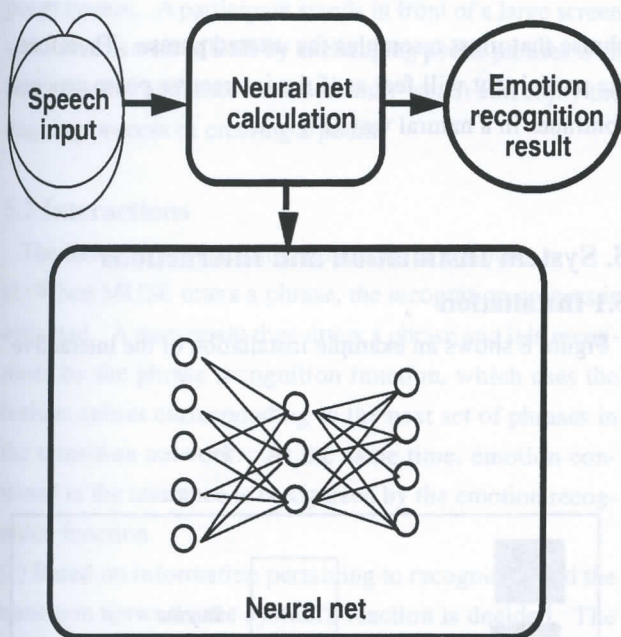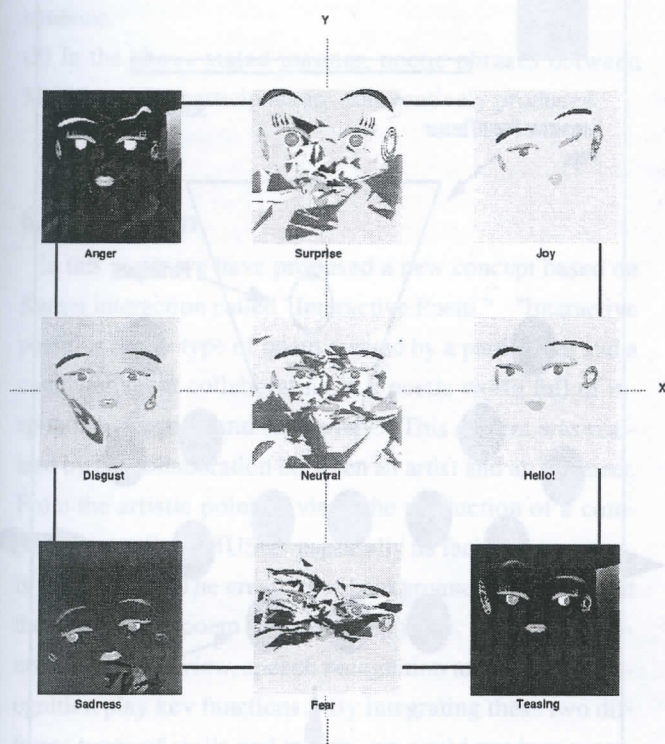

Fig. 3 Phrase recognition

Fig. 4 Emotion recognition

Fig. 5 Muse's emotional expression

realized (Fig. 4).

Reaction of the computer agent to utterances of a participant is expressed through her speech and by images. In the speech output unit, speech data for each phrase to be uttered by the computer agent is digitally stored and generated when necessary.

The computer graphics generation unit controls image reaction of the computer agent. Image reaction consists of two kinds of images: facial expressions for the computer agent "MUSE" and various scenes. The facial expressions of MUSE express her reactions to the emotional state of a participant. These images are represented by keyframe animations, each of which corresponds to the eight emotions (Fig. 5). To express the atmosphere of the interactive
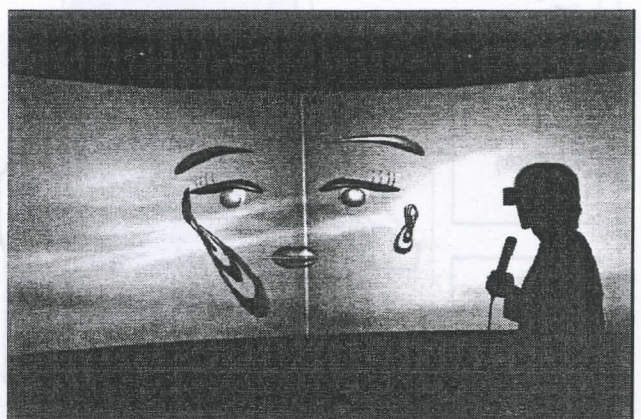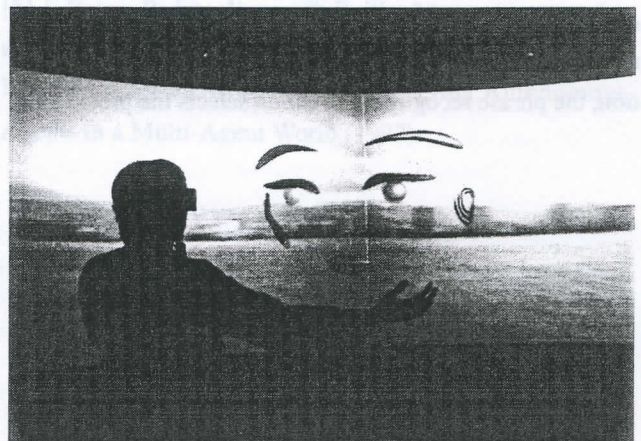
Fig. 6 Example of interaction between Muse and the participant

poem, several kinds of scenes are digitally stored. Each scene image corresponds to a group of states in the transition network, and each correspondence is carefully determined in advance (Fig. 6).

4. Hardware Configuration

This mainly consists of several workstations and a PC: a workstation for computer graphics generation, a workstation for both system control and phrase recognition, a workstation for emotion recognition, and a PC for speech output. For the participant's convenience, optional phrases that may be uttered following an utterance of MUSE appear on the display. The participant can choose one of these phrases based on their feelings and sensitivity, or they can create their own poetic phrase. Regardless, the emotion recognition function can produce a result. In addition, the phrase recognition function selects the preexisting

phrase that most resembles the uttered phrase. Therefore, the participant will feel as if the interactive poem process continues in a natural way (Fig. 7).

5. System Installation and Interactions

5.1 Installation
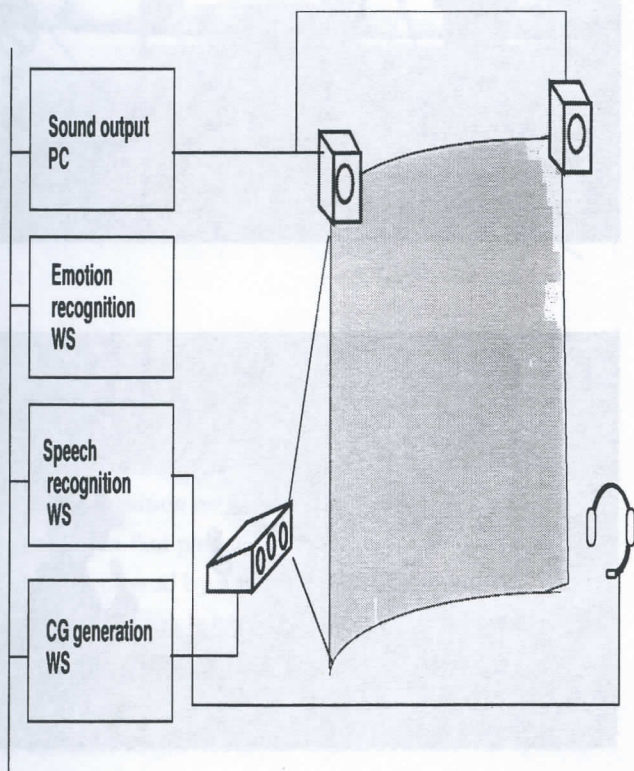
Figure 8 shows an example installation of the interactive

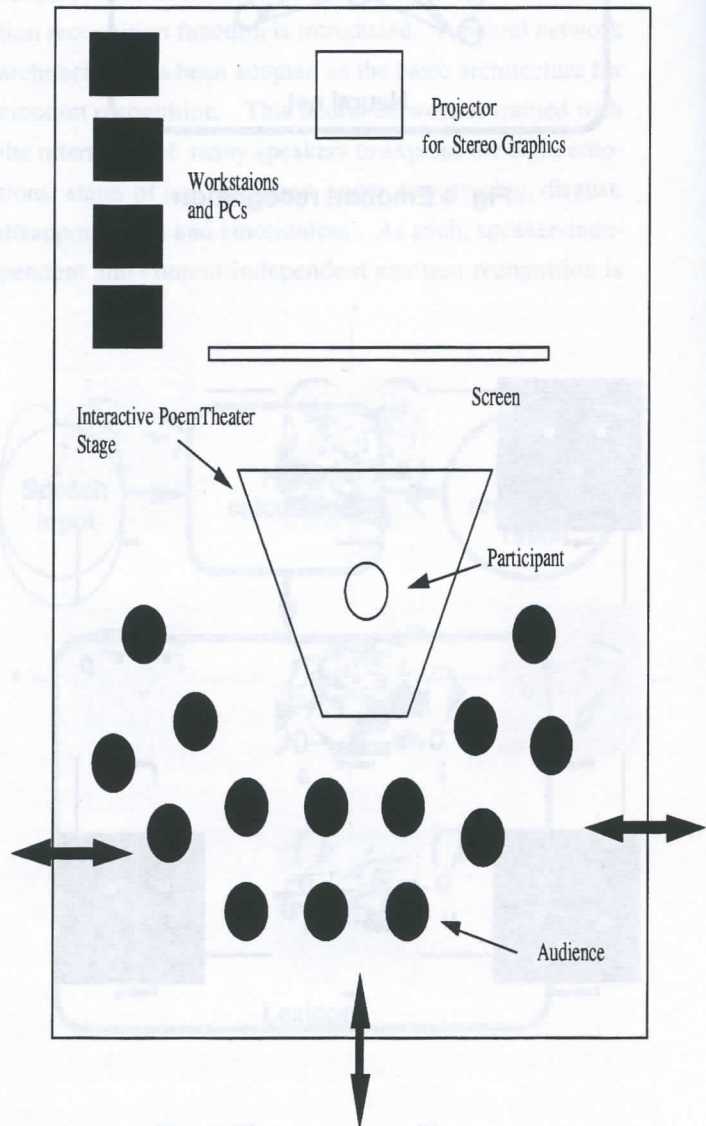**Fig. 7 Interactive Poem hardware configuration**

Fig. 8 Installation

poem system. A participant stands in front of a large screen and interacts with MUSE by exchanging poetic phrases with her. Several audience members may watch and enjoy the ongoing process of creating a poem.

5.2 Interactions

The interaction mechanism operates as follows.

(1) When MUSE utters a phrase, the recognition process is activated. A participant then utters a phrase and it is recognized by the phrase recognition function, which uses the lexicon subset corresponding to the next set of phrases in the transition network. At the same time, emotion contained in the utterance is recognized by the emotion recognition function.

(2) Based on information pertaining to recognition and the transition network, the system's reaction is decided. The facial expression of MUSE changes according to the results of emotion recognition, and the phrase MUSE utters is based on the results of phrase recognition and the transition network. The background scene changes as the transitions continue.

(3) In the above stated manner, poetic phrases between MUSE and the participant are consecutively produced.

6. Conclusion

In this paper we have proposed a new concept based on Kansei interaction called "Interactive Poem." "Interactive poem" is a new type of poem created by a participant and a computer agent collaborating in a poetic world full of inspiration, emotion and sensitivity. This system was realized by the collaboration between an artist and an engineer. From the artistic point of view, the production of a computer poet called "MUSE," especially its facial expressions, is a key issue. The creation of background images that fit the mood of the poem is another key issue. From the technology point of view, speech recognition and emotion recognition play key functions. By integrating these two different types of skills and talents, we could produce a new system that can be considered both a new interactive art medium and a new Kansei-based interaction system.

References

[1] N. Tosa et al., "Neuro-Character," AAAI'94 Workshop, AI and A-Life and Entertainment (1994).

[2] N.Tosa and R. Nakatsu, "Life-like Communication Agent--Emotion Sensing Character 'MIC' and Feeling Session Character 'MUSE'," Proceedings of the International Conference on Multimedia Computing and Systems, pp.12-19(1996).

[3] P. Maes, T. Darrell, B. Blumberg, and A. Pentland, "The ALIVE system: Full-body interaction with autonomous agents," Proc. of the Computer Animation'95 Conference (1995).

[4] K. Perlin, "Real time responsive animation with personality," IEEE Transactions on Visualization and Computer Graphics, Vol.1, No.1 pp.5-15 (1995).

[5] J. Bates, B. Loyall, and S. Reilly, "An architecture for action, emotion, and social behavior," Proceedings of the Fourth European Workshop on Modeling Autonomous Agents in a Multi-Agent World (1992).

This volume includes the papers presented at the *AIMI International Workshop on Kansei – The Technology of Emotion* (Genova, 3-4 October 1997), organized by AIMI (Associazione di Informatica Musicale Italiana), Laboratorio di Informatica Musicale at DIST-University of Genova, and Teatro dell'Opera di Genova Carlo Felice. Main topics include expressivity in virtual musical instruments; adaptive hyper-instruments; computational models of artificial emotions; sensibility and sensual intelligence; understanding musical emotions; agent models integrating rational and emotional components; relations between music and gesture languages; on-stage real-time multimodal environments; interactive dance/music systems: dance interpretation and integration with music languages; adaptive listening; modeling of expressive performance.
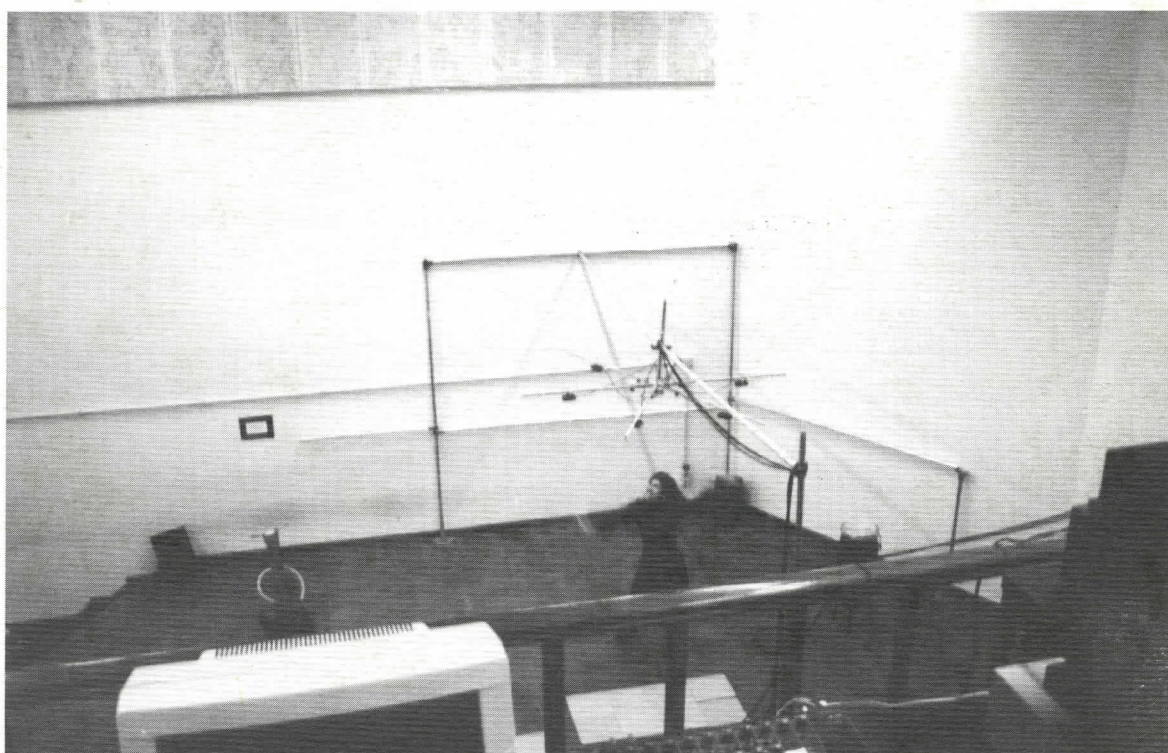
Teatro dell'Opera di Genova Carlo Felice: The Laboratorio di Informatica Musicale of DIST-University of Genova