An Experimental Multimodal Disputation System

Katsumi Nitta * and **Osamu Hasegawa** and **Tomoyoshi Akiba**
Toshihiro Kamishima and Takio Kurita and Satoru Hayamizu and Katsunobu Itoh
(Electrotechnical Laboratory)
Mitsuru Ishizuka and Hiroshi Dohi
(Tokyo University)
Manabu Okumura
(Japan Advanced Institute of Science and Technology)

Abstract

An experimental multimodal disputation system, Mr.Bengo, is a knowledge based system with multimodal user interfaces such as face recognition, face generation, speech recognition, speech generation and a WWW browser. Mr. Bengo deals with three agents - a prosecution, an defense attorney and the judge. The prosecution and the attorney dispute about the interpretation of legal rules. The user controls the defense attorney to dispute with the prosecution. After the disputation finishes, the judge decides the winner.

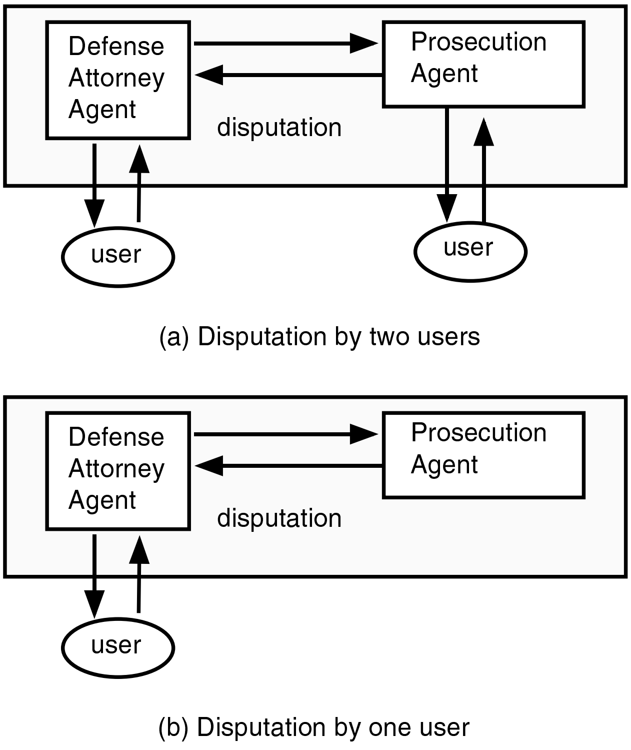
As the disputation is a two agent game, to predict the opponent's move is important to win the game. During disputation of the Mr. Bengo, the face of each agent changes according to the status of disputation. Using the face information, the user knows if the prosecution finds the counter argument or not, which helps him to search the good next move.

1 Introduction

As legal knowledge is given by the statutory rules, it seems easier to develop legal knowledge bases. However, there are several difficulties in statutory rules. One is that some legal rules conflict each other. And the other is that some legal rules are given by ambiguous predicates. Therefore, when they are applied to actual cases, their meaning must be clarified by interpreting them. However, as there is no definite way for interpretation, in the legal court, the prosecution (or the plaintiff) and the defense dispute by insisting different interpretations.

On the analysis of legal disputes, we have developed a legal reasoning system, the New HELIC-II [2] in the Fifth Generation Computer Systems Follow-up Project. This system consists of two software agents (a prosecution agent and a defense agent), and two users dispute through these agents (Fig. 1(a)). If there is only one user, he can dispute with computer. In such case, the prosecution agent behaves autonomously by using a predefined disputation strategy (Fig. 1(b)).

The New HELIC-II has a graphical user interface, and the users have to indicate the next moves to their agents by selecting commands in their windows. Though computer scientists are accustomed to the multiple windows interface, it is troublesome for lawyers to use it.



(a) Disputation by two users

(b) Disputation by one user

Figure 1: Disputation of the New HELIC-II

Apart from the research of the New HELIC-II, we have developed multimodal interactive technologies in the Real World Computing Systems Project, Tokyo University and Japan Advanced Institute of Science and Technology. The objective of this research is to realize natural interactive systems by combining software agent technologies, image processing technologies, speech signal processing technologies and natural language understanding technologies. We have shown the effects of combining these technologies by developing several experimental systems.

Mr. Bengo is developed by combining these technologies. It is our first trial which unifies the complex knowledge based system and the agent oriented multimodal information processing technologies.

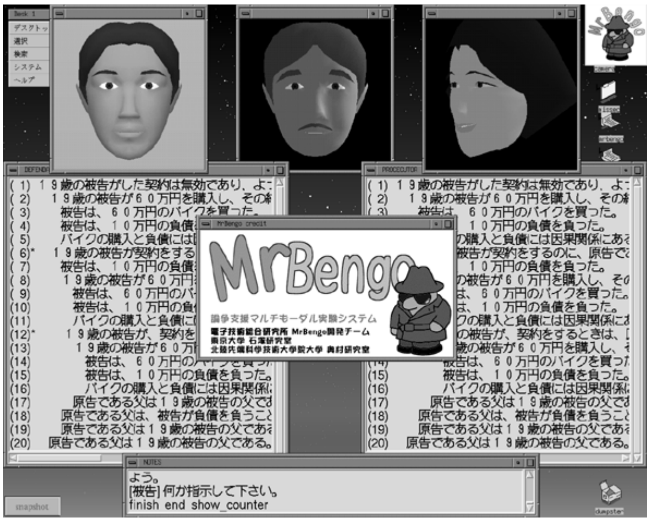
The objective of this research is to show the effectiveness of multimodal user interface in the disputation system. As the disputation is a two agent game, the face representation of the agent affects the strategy of the opponent agent to select the next move.

In Section Two, we will show the architecture and the functions of Mr.Bengo. In Section Three, submodules of Mr.Bengo are introduced, and in Section Four, we will discuss the effectiveness of face representation during disputation.

2 Overview of Mr.Bengo

Mr.Bengo is implemented on the SGI workstation INDYR4400 200MHz with a TV camera (IndyCam), a microphone, a DAT player and a speech synthesizer (Shaberinbou).

When Mr.Bengo system starts, it observes the user's face in the IndyCam and compares it to face images in the database. If Mr. Bengo succeeds in recognizing the user, three faces appear in the screen. These faces corresponds to three agents - prosecution, the defense attorney and the judge (Fig.2).



(1) １９歳の被告がした契約は無効であり、よ
(2) １９歳の被告が６０万円を購入し、その
(3) 被告は、６０万円のバイクを買った。
(4) 被告は、１０万円の負債
(5) バイクの購入と負債には
(6)* １９歳の被告が契約をする
(7) 被告は、１０万円の負債
(8) １９歳の被告が６０万円
(9) 被告は、６０万円のバ
(10) 被告は、１０万円の負
(11) バイクの購入と負債に
(12)* １９歳の被告が、契約を
(13) １９歳の被告が６０万
(14) 被告は、６０万円の
(15) 被告は、１０万円の負債を負った。
(16) バイクの購入と負債には因果関係
(17) 原告である父は１９歳の被告の父で
(18) 原告である父は、被告が負債を負うこと
(19) 原告である父は１９歳の被告の父であ
(20) 原告である父は１９歳の被告の父である。

(1) １９歳の被告がした契約は無効であり、よ
(2) １９歳の被告が６０万円を購入し、その
(3) 被告は、６０万円のバイクを買った。
１０万円の負債を負った。
の購入と負債には因果関係にあ
被告が契約をするのに、原告で
１０万円の負債を負った。
の被告が６０万円を購入し、そ
は、６０万円のバイクを買った。
は、１０万円の負債を負った。
クの購入と負債には因果関係に
の被告が、契約をするときは、
歳の被告が６０万円を購入し、
被告は、６０万円のバイクを買っ
(15) 被告は、１０万円の負債を負った。
(16) バイクの購入と負債には因果関係
(17) 原告である父は１９歳の被告の父で
(18) 原告である父は、被告が負債を負うこと
(19) 原告である父は１９歳の被告の父であ
(20) 原告である父は１９歳の被告の父である。

MrBengo
論争支援マルチモーダル実験システム
電子技術総合研究所 MrBengo開発チーム
東京大学 石塚研究室
北陸先端科学技術大学院大学 奥村研究室

よう。
[被告] 何か指示して下さい。
finish end show_counter

Figure 2: Initial stage of Mr.Bengo

The user disputes with the prosecution as following example.

(1) At first, the prosecution agent presents the initial argument to the defense attorney agent and the user.

(2) The user finds some issue point in the argument, and tells the defense attorney to make a list of counter arguments for the issue point. The user selects the best counter argument in the list and tell the defense attorney to present it to the prosecution.

(3) The prosecution compares his argument and the defense's counter argument, and insists that his original argument has priority over the counter argument.

(4) The user consults to judicial precedents through a WWW browser to get materials to insist the counter argument has the priority. However, as he fails to find the proper old case, he asks the defense attorney to withdraw the counter argument.

(5) The defense attorney informs the prosecution to withdraw the counter argument. Then, the judge agent announces that the prosecution wins.

During the disputation, the user communicates with agents only by speech, and he needs not to use a keyboard and a mouse. In addition to it, as the face image of each agent changes according to internal state of the agent. If the prosecution smiles, the user estimates that it is difficult to find a good counter argument because the prosecution thinks he is sure to win. If the prosecution looks sad, the user may win because the prosecution noticed that there is a good counter argument.

3 Modules of Mr.Bengo

Mr. Bengo consists of several modules (Fig. 3). The basic modules are a disputation, a face recognition, a speech recognition, a face image synthesis, a speech synthesis and a WWW browser. They are connected by three managers.

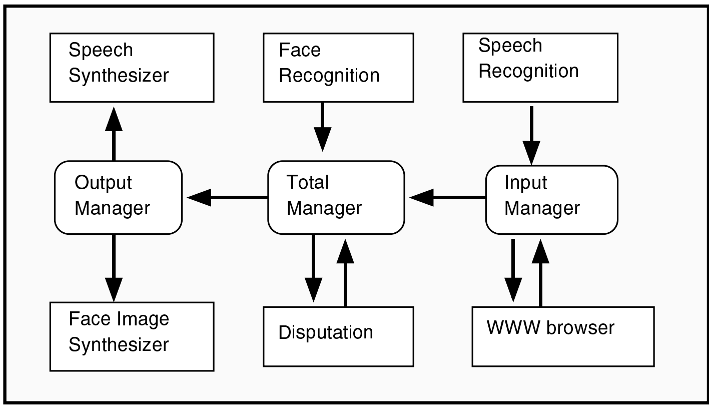


Figure 3: Architecture of Mr.Bengo

In this section, we will introduce the function of each module.

3.1 Disputation module

The disputation module defines the functions of three agents. The judge agent has two functions - 'making arguments' and 'comparing arguments'. The defense attorney agent has three functions - 'making arguments', 'comparing arguments', and 'disputation moves'. The prosecution agent has four functions - 'making arguments', 'comparing arguments', 'disputation moves' and 'disputation procedure' (Fig.4). Each agent has a knowledge base which contains statutory rules, the judicial precedents and common sense knowledge, and priority of standpoints. *Standpoints* are criteria which are used to compare the priority of conflicting arguments. 'Protecting freedom of press', 'protecting human rights' and 'protecting social morals' are examples of standpoints of the criminal law.

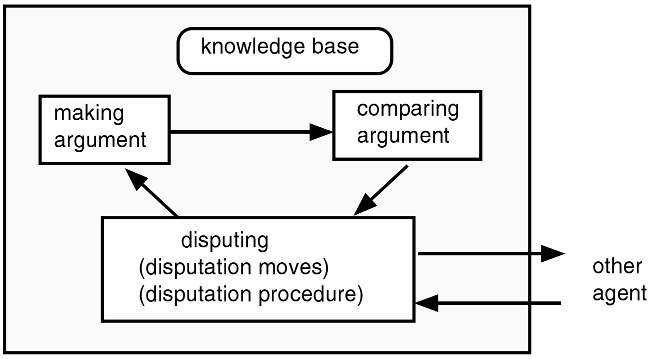


Figure 4: The disputation module

(i) Making argument:
The 'making argument' submodule makes arguments for a given conclusion (goal) by refering to rules. An *argument* is a minimal set of grounded rules which draws the goal. A *counter argument* is an argument whose conclusion is a complement of the intermediate goal of the original argument.

Following is an example of an argument and a counterargument. *ArgA* is an argument whose conclusion is $p(a)$, and *ArgC* is an counterargumment for *ArgA* because its final conclusion $\neg q(a)$ is a complement of $q(a)$ which is an intermediate goal in *ArgA*.

$$\begin{array}{lll} \text{Ex.)} & \text{ArgA } \{ & r1 : p(a) \leftarrow q(a), s(a)., \\ & & r2 : q(a) \leftarrow r(a)., \\ & & f1 : r(a)., \quad f2 : s(a).\ \}. \\ & \text{ArgC } \{ & r3 : \neg q(a) \leftarrow t(a)., \\ & & f3 : t(a).\ \}. \end{array}$$

(ii) Comparing argument:
The 'comparing argument' submodule compares an argument and a counter argument, and decides which has priority. The priority of an argument is defined by the priority of rules included in each argument. Furthermore, the priority of a rule is defined by the priority of the standpoint of the rule. For above example, top rules of *ArgA* and *ArgC* are $r2$ and $r3$. Let $r2$ and $r3$ contain standpoints u and v respectively, and if u has priority over v, then *ArgA* has priority over *ArgC*.

Generally, as the prosecution and the defense have different opinions about priority of standpoints, it can happen that the prosecution insists *ArgA* has priority and the defense insists *ArgC* has priority. In such case, the judge makes the final decision according to the judge's opinion.

(iii) Disputation moves:
The 'disputation moves' submodule prepares following disputation moves which the prosecution agent and the defense attorney agent can take.

(1) **claim(Argument):** An agent presents an argument to the opposite agent.

(2) **show_counter(Goal):** An agent makes a list of arguments for a given goal by using the 'making argument' submodule.

(3) **justify(Argument1, Argument2):** An agent makes a hypothesis of priority of standpoints in order that *Argument1* has priority over *Argument2*.

(4) **priority(Argument1,Argument2):** An agent calculates priority between Argument1 and Argument2.

(5) **notify(Priority):** An agent notifies the opposite agent that an argument has priority over a counter argument or not.

(6) **cancel(Argument):** An agent withdraws an argument.

(7) **finish:** An agent tells the opposite argument that he wishes to finish the disputation.

The user of Mr.Bengo controls the defense attorney agent by indicating one of above moves. For example, when the prosecutor agent presents him an argument, the user selects an issue point, and asks the defense attorney agent to make a list of counter arguments for it by 'show _counter' command. If he finds a good counter argument in the list, he presents it to the prosecution agent by 'claim' command. However, if he doesn't find a proper counter argument, he may finish the disputation by 'finish' command.

(iv) Disputation procedures (Disputation strategy):
While the user of Mr.Bengo controls the defense attorney agent, the prosecution agent behaves autonomously by using the disputation procedures.

The disputation procedure manages a progress of a disputation as a regular automaton and gives an algorithm to select the next move for each state of the automaton. Following is an example of the disputation procedure.

- $() \rightarrow$ (*show_counter G List*)
- (*show_counter G List*) $\rightarrow$ (*claim A nil*)

- (*show_counter G List*) → (*finish A nil*)
- (*cancel* (*claim B* (*issue C i*))) → (*issue C i*)

(v) An example of disputation:

Following is an example of a disputation of the criminal case. 'P', 'D', and 'A' mean an prosecution agent, a defense (the user) and a defense attorney, respectively.

- (P→D:) "The defense should be punished by a crime of death by negligence because he caused traffic accident after he drank the beer and the victim died of it."
- (A→D) "What shall I do?"
- (D→A) "I don't agree that the victim died of the accident. Please try to make counter arguments for it."
- (A→D) "First counter argument is that the victim didn't died of the accident because he was not injured so much by the accident, and he had been suffered from high blood pressure for a long time."
- (D→A) " Show it to the prosecution."
- (D→P) "The victim didn't died of the accident because he was not injured so much by the accident, and he had been suffered from high blood pressure for a long time."
- (P→D) "The victim had not been suffered from high blood pressure so much because he had not taken medicine for it."
- .....

3.2 Face recognition module

The 'face recognition' module observes the human face in the TV camera, compares it to pre-registered face images and recognizes who he is.

To recognize faces, this module extracts the following higher order local self correlation feature from training face images.

$$x^N(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_N) = \int f(\boldsymbol{r}) f(\boldsymbol{r} + \boldsymbol{a}_1) \cdots f(\boldsymbol{r} + \boldsymbol{a}_N) d\boldsymbol{r} \tag{1}$$

Here, $f(\boldsymbol{r})$ is a image in the TV, $(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_N)$ is displacement.

The initial image feature based on this function is combined by the following function.

$$y_j = \sum_{i=1}^{M} a_{ij} x_i \quad (j = 1, \ldots, N) \quad \boldsymbol{y} = A^T \boldsymbol{x} \tag{2}$$

Here, $\{x_i | i = 1, \ldots, M\}$ is an initial image feature, $\{a_{ij}\}$ is a coefficient, M is a number of the initial image feature, and A^T is a transposed matrix of $A = [a_{ij}]$.

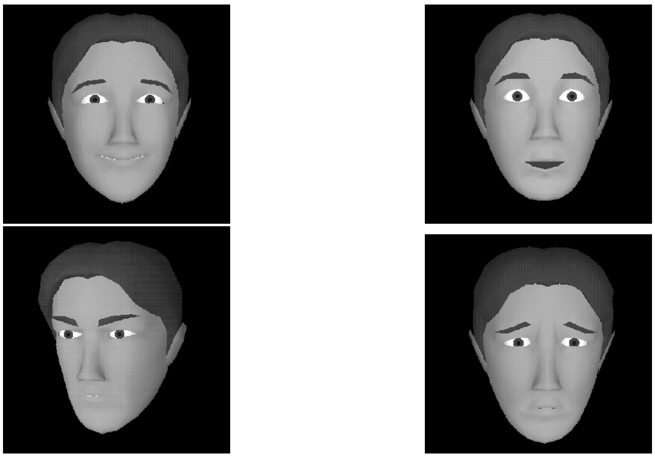


Figure 5: Example of expressions of agents

If a set of samples of K classes of initial feature vector $\boldsymbol{x}$ is given as $C_k = \{\boldsymbol{x}\}$ $(k = 1, \ldots, K)$, the coefficient matrix A for linear discriminant analysis is obtained by solving following equation.

$$\Sigma_B A = \Sigma_W A \Lambda, \quad A^T \Sigma_W A = I_N \tag{3}$$

Here, N is a dimension of discriminant space, and $N \leq min(K-1, m)$. Σ_B and Σ_W are inter-class covariance matrix and a within-class covariance matrix.

When a new unknown image is given, its class is estimated by calculating $\bar{\boldsymbol{y}}_k = A^T \bar{\boldsymbol{x}}_k$. Here, $\boldsymbol{x}$ is an initial feature of the new image.

3.3 Face image synthesizing module

The 'face image synthesizing' module generates face images of three agents.

Each face image model consists of 800 vertex and 1500 poligons which have three dimensional coordinates. Each face has five expressions such as a normal image, a angry image, a sad image, a delight image and a surprising image, and has several actions such as nodding and winking (Fig. 5).

Drawing faces of agents are implemented by using C++, OpenGL and GLUT. Following *command, mface, agent, obj* are names of programs included in this module.

(1) *command*: It receives all drawing commands through the socket.
(2) *mface*: It manages action of each agent based on the received command, and sends drawing commands to each agent.
(3) *agent*: It analyzes commands from *mface*, and sends drawing commands to each *obj*.
(4) *obj*: It recognizes the status of an agent and the positions of the point of view and the light, and requires of drawing to OpenGL.
(5) OpenGL: It draws the face of an agent.

(6) GLUT: It manages the window to which the OpenGL draws.
(7) X: It presents windows to three agents, and displays them on the screen.

3.4 Speech recognition module

This module consists of 'Speech feature extraction submodule' and 'speech management submodule'.

'Speech feature extraction submodule' analyzes speech signal from a microphone, and generates a recognized sentence as follows.

It recognizes the signal part from the noise part, extracts possible sequence of phonetics, selects the best sequence of words by refering to a word dictionary and Japanese grammar.

'Speech management submodule' receives the recognized Japanese sentence, make up with the context information, and generates a dispute command.

3.5 WWW browser

The 'WWW browser' is controlled by the speech. It consists of an interface library written in C and the modified NCSA Mosaic which is controlled by outside processes.

The interface library includes following three functions which are used to select WWW links.

1. Selection by anchor letters
   By speaking the anchor letters in the Mosaic screen, WWW links are retrieved.
2. Selection by index number
   When a new WWW page is opened,anchor list is presented to the users.
3. Selection by URL

3.6 Speech synthesizing module

We use the speech synthesizing module which is installed into the speech synthesizer (Shaberinbou).

3.7 Managers

Basic modules which are introduced in previous sections are combined by three managers - a total manager, input manager and output manager.

The 'total manager' manages communication among each module. As each basic module has different I/O protocols, the total manager adapts them by changing the form of some module's output to form of another module's input.

The 'output manager' controls two modules - the face image synthesizing and the speech synthesizing. When each agent receives an argument from the opposite, if his argument has priority, he smiles. However, if opposite's argument has priority, his face becomes sad. Furthermore, if he found that it is difficult to find the effective counter argument, his face becomes angry.

The 'input manager' controls the 'speech recognition' module. While some input commands are sent to the disputation module, and other commands are sent to the WWW browser.

4 Face information and disputation

In this section, we will show that the expression of face presents important information to the user.

As presented in the previous section, disputation is conducted by exchanging disputation moves each other. As there are several candidates of issue points and several candidates of counter arguments, the user has to select one for each disputation step. If some counter argument contains a conflict with the arguments which the user already presented previously, the user cannot present it to the opponent. Therefore, if the user's disputation strategy is inappropriate, he may loose the disputation.

In our disputation model, the prosecution agent and the defense attorney agent has the different rule base and the different priority of standpoints. Though both agents don't know the opposite's knowledge, if they estimates it, they can select the better move. They estimate opposite's knowledge by observing disputation moves which the opposite has taken. For example, if the prosecution insists some argument has prioriry over the other argument, we can estimate that some standpoint has priority over the other standpoint in his opinion. The estimation is realized by generating hypothesis by using abductive reasoning technique. However, generally, the estimation is not easy because there are lots of possible hypotheses and there is no definite way to select the best one.

In Mr.Bengo, the estimation of opposite's knowledge becomes easier because the user can observe the face of the prosecution agent changes. When the user presents the counter argument to the prosecution, the user understands how the prosecution evaluates it because if the counter argument has priority, his face becomes sad state. Therefore, by observing representation of the face of the prosecution agent, the quality of estimation of opposite's knowledge is improved.

5 Conclusion

We presented the overview of an experimental multimodal disputation system, Mr. Bengo. This system shows that the natural user interface is realized by using multimodal information processing technologies. In addition to it, we showed the representation of face gives useful information for selecting the next disputation move because the defense can estimate the prosecution notices a candidate of effective counter argument or not. As our current disputation model is a kind of two agent games such as Chess and Shogi, our research may be applicable to the instructing system for these games.

We have been expanding Mr. Bengo to treat more general negotiation which includes not only disputation but reconciling. In the negotiation system, the role of the face representation becomes more important.

References

[1] [Nitta *et. al.*, 1996] Nitta, et. al. : An experimental multimodal disputation system (in Japanese), In proc. Tsukuba Software Symposium 96, Tsukuba, 1996.

[2] [Nitta *et al.*, 1995] Nitta, et. al. : New HELIC-II: a software tool for legal reasoning, In proc. Int. Joint. Conf. on AI and Law, Maryland, 1995.

[3] [Itou *et al.*, 1990] Itou, et al., Continuous speech recognition using extended LR syntax analysis (in Japanese), Report of IEICE, Vol.SP90, No.74, pp.49-56, 1990.

[4] [AIba *et al.*, 1996] AIba,et.al., Research of semantics analysis of speech conversation (in Japanese), Research Report of ETL, 1996.

[5] [Kurita *et al.*, 1992] Kurita T., Otsu N. and Sato T., "A face recognition method using higher order local autocorrelation and multivariate analysis," Proc. of 11th ICPR, The Hague, Vol.II, pp. 213-216, 1992.

[6] [Hasegawa *et al.*, 1995] O.Hasegawa et al.:"Active Agent Oriented Multimodal Interface System", Proc. AAAI International Joint Conference on Artificial Intelligence 95, Vol.1, pp.82-87 1995

[7] [Dohi and Ishizuka, 1996] H.Dohi and M.Ishizuka: A Visual Software Agent connected with the WWW/Mosaic, *Int'l Symp. Multimedia Systems (MMJ '96)*, pp. 392–397, 1996.