CORBA Based Distributed Computing Model for Multimodal Speech Recognition*

Meng Shan Zhou Yanlei Zhang Youwei
Wuyi University, Institute of Information Science, Jiangmen, 529020 Guangdong, China
Beijing University of Aeronautics and Astronautics, Dept. of EE, 100083 Beijing, China
E-mail: shmeng@163.net

ABSTRACT

In this paper, we presented our work in progress on computing infrastructure of HMNI (Human-Machine Natural Interaction) research project. In multimodal speech recognition system, various signals and huge amount of data are collected and analyzed simultaneously. To make this kind of task realizable, we developed a practical distributed computing model complying with CORBA (Common Object Request Broker Architecture).

1. INTRODUCTION

Speech recognition has evolved into the main method of the next generation human-machine interaction. Unlike man to man interaction, the pure speech recognition system forces user to interact with microphone in relatively quiet environment. The communication between human and machine is quite inefficient. Its quality is easily degraded by various kinds of background noise.

To make the human-machine interaction more natural, we must employ the visual modalities in HMNI system just like the ordinary human communication process. There are many kinds of visual cues that can help to increase the accuracy and robustness of speech recognition, such as eye/lip movement, gesture, and handwriting etc. The key problem here is how to effectively combine these visual modalities with speech information to achieve natural human-machine interfaces.

As we know, visual information processing is computation intensive. HMNI must be realized as heterogeneous and distributed computing system. Figure 1 shows the diagram of our HMNI design. Details of HMNI are presented by Meng [1].

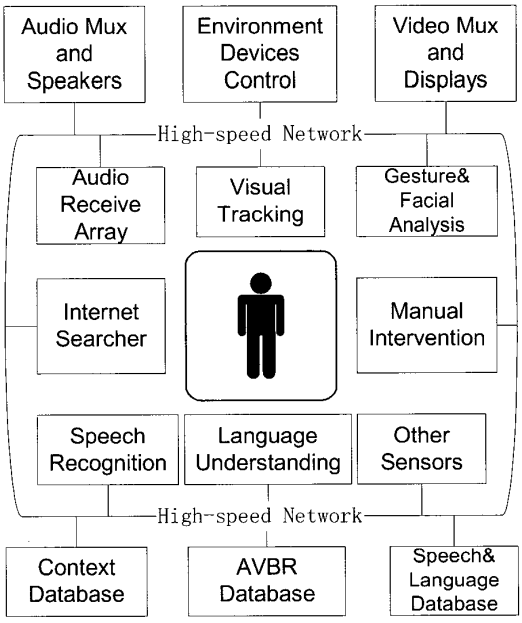

Figure 1: Basic System Structure of MMNI

A similar model was presented by Wei [2]. It is called Distributed Speech Recognition (DSR) model. The DSR architecture is basically a client/server realization of distributed computing. The client devices collect the speech signal and transfer extracted features, such as MFCC vector, to the DSR server, then wait for the server's recognition result to give the appropriate feedback to users. This model can alleviate the burden of voice enabled client and the system optimization can be mainly applied to the distributed recognition server [2]. In our work for the HMNI's distributed computing environment, we found that it's too complex to realize if we use the client/server model to develop and deploy every

* This work was supported by NSF of Guangdong, China

component of HMNI. For such a complex system, we must use the object-oriented method to realize each component during the development phase. Beyond the Client/Server model, an object can be server and client simultaneously. The main targets of our design are object development separation, better coordination and objects' communication decoupling. Our solution architecture is CORBA.

In this paper, concepts of CORBA are discussed in section 2. Our work on distributed computing model is presented in section 3. Some performance considerations are proposed in section 4.

2. Fundamentals of CORBA

CORBA is an open, standard solution for distributed object systems. To make it simple, Figure 2 illustrates the client/server view of CORBA. OBJECT 1 and OBJECT 2 are objects in distributed computing system, for example, HMNI system. According to the OO (Object Oriented) method, objects are encapsulated with their own attributes and methods. Each of them is full-fledged component, which acts as server as well as client. With the support of distributed computing environment, they can transparently make method invocation on other objects and access their attributes.

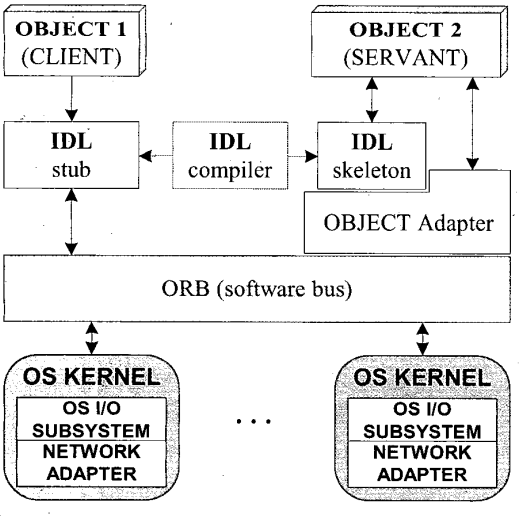

Figure 2: Client/Server View of CORBA

ORB (Object Request Broker) is the heart of the distributed system. It is a software bus through which all the objects are linked together. ORB abstracts the platform specific OS (Operating System) I/O subsystem and network connection for objects above. There are a series of CORBA services to support interoperation between distributed objects. For example, naming service can help an object to get another object's reference, via this reference the object can request the other object's service. CORBA-compliant ORBs can communicate with each other via GIOP (General Inter-ORB Protocol).

IDL (Interface Definition Language) is defined by OMG (Object Management Group) to describe an object's outside behavior. An Object is an instance of an IDL defined interface. The IDL compiler processes the IDL file and generates stub for client side and skeleton for server side. There are IDL language mappings defined for several programming languages, such as C, C++ and Java. We can use these programming languages develop our own objects. They inter-operate with other objects with support of GIOP, which is usually implemented as IIOP (Internet Inter-ORB Protocol), without caring about their programming language [3].

Stubs act as "glue" between the client object and the ORB. Stubs marshal the client objects' request into common message-level representation to ensure that ORB understands how to deliver it to the server object [4]. Stub also decodes the returned result or exception for the client object.

Skeletons act as "glue" between the server object and the ORB. Skeletons demarshal the message-level request from client object back into conventional method call with typed parameters that can be understood by servant [4]. It also marshals the result or exception produced by servant into message-level form, and then delivered to client by ORB.

OA (Object Adapter) is crucial to the distributed computing system. It is the proxy between the server host and ORB. Tasks carried out by OA include object activation/deactivation, requested operation dispatching and object references generation etc [5]. A host machine in distributed system usually contains multiple instances of objects. So the resource management and task schedule at server host will greatly influence the performance of distributed

computing.

With the characters discussed above, we can see that CORBA supplies a uniform running environment for objects that are developed in different languages and different platforms. It's great for teamwork. We can develop all the objects separately. As long as they implement interfaces predefined by IDL files, they will coordinate with each other fluently, and communicate with others transparently.

3. HMNI COMPUTING MODEL

For HMNI, multimedia stream that includes huge amounts of data continuously flow through the distributed processors, so the real-time capability is necessary. CORBA is built above the classic network protocol, so there are doubts about if the performance requirement can be met with this kind of architecture. Actually, its inherent drawback on performance can be disregarded when comparing with the advantage it brings to us. Especially, the speed of processor and bandwidth of network have dramatically increased and they will keep on increasing.

Our HMNI is at its infancy phase of development. It involves many technology domains, such as signal processing, network management, multiple database management and Internet access etc. There will be lots of objects coexist and cooperate. We had developed several toolkits used for distributed object development. The development combination is Borland Visibroker and Java.

To realize distributed objects enumeration, we developed a consulting toolkit named as Enumerator. Figure 3 shows the working window of it.

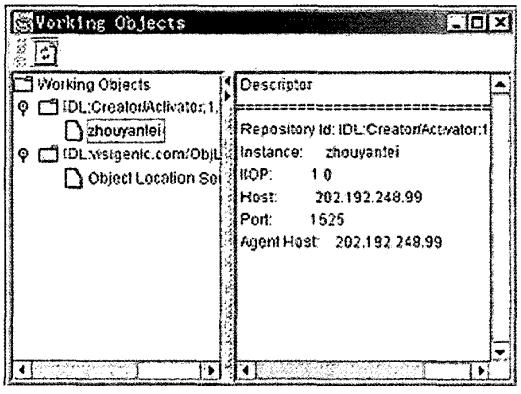

Figure 3: Objects Enumerator

The enumerator collects information of objects running in the distributed system. Enumerator can reveal each object's information in detail, such as its host IP address, port number and its reference ID.

To reduce the development complexity, we developed a central configuration toolkit. Figure 4 illustrates its appearance.

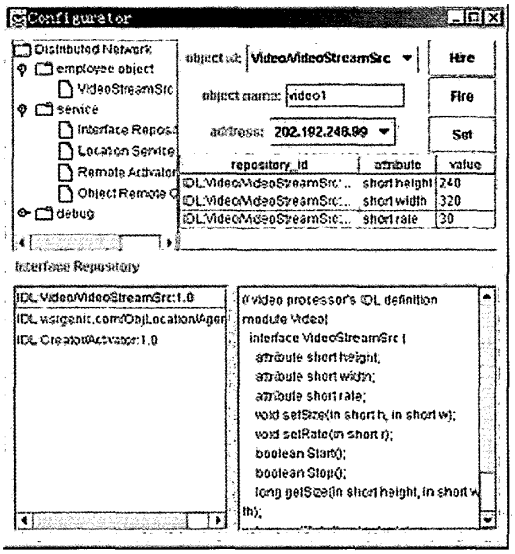

Figure 4: Configuration Toolkit

Through the configuration toolkit, we can set specific object's initial parameters, remote activate/deactivate an object on specific host and check each object's IDL interface definition.

A testing toolkit was developed to help developer to test under developing objects' function and properties. Figure 5 shows the Tester's initial window in which there is a list of available objects. If an object is selected from the list, there will popup a window containing the object's available methods and properties. Figure 6 shows a running object's testing process. Test can be carried on any method or property of the object.

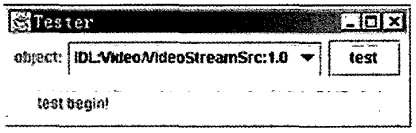

Figure 5: Tester's initial window

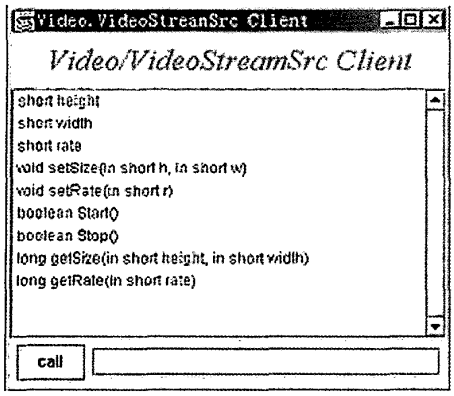

Figure 6: Testing window

All the toolkits above are developed according to the CORBA specification. Each of them is an object in HMNI system. They make the distributed developing process more scalable, configurable and adaptable.

4.PERFORMANCE CONSIDERATION

4.1 Quality of Service (QoS) consideration

QoS describes activities and technologies designed to improve and control communication-oriented resource management for applications and systems. The OMG messaging specification defines a QoS framework [6]. HMNI system is a time stringent distributed system. The network throughput, operation priority and transit delay must be controlled by the QoS framework.

4.2 Choose appropriate ORBs

There are many available ORB products in the market and free software world. For the time stringent parts of HMNI, we must choose the real-time ORBs to meet the performance requirement, for instance, TAO[1] — an open source real-time ORB developed by Washington University. For the less time stringent parts of HMNI, we can use some commercial products with great flexibility and portability. Such as Visibroker and Orbix's ORBs.

5. CONCLUSION

Multimodal speech recognition system, such as HMNI, is a quite complex distributed computing system. The development of such a heterogeneous system must be well planned from the beginning. Our CORBA base distributed computing model can solve this problem. Our future work will focus on audio and video streaming and real-time distributed signal processing.

REFERENCE

[1] Meng Shan, et. al. A new basic system structures for Man-Machine Natural Interaction and its kernel techniques. In Proceedings of ICSSP'2000, IEEE Press, 2000, pp. 1031-1034.

[2] WeiQi, et. al, The study on distributed speech recognition system. In Proceedings of ICASS'2000, IEEE Press, CDROM: 135_169.PDF.

[3] Samuel Bayer, Embedding speech in web interfaces. In Proceedings of ICSLP'1996, pp. 1684-1686.

[4] Irfan Pyarali, et. al. Applying optimazation principle patterns to real-time ORBs. http://www.cs.wustl.edu/~schmidt/PDF/IEEE-demuxing.pdf.

[5] Douglas C. Schmidt, et. al. Object Adapters: Concepts and Terminology. C++ Report, SIGS, Vol. 9, No 11. November/December, 1997.

[6] Douglas C. Schmidt, et. al. An Overview of the OMG CORBA Messaging Quality of Service (QoS) Framework. C++ Report magazine, SIGS, Vol. 12, No 3, March, 2000.

[1] web site: www.cs.wustl.edu/~schmidt/TAO.html