

Comparison of the virtual environment Implementation with Services of speaks

Eduardo Filgueiras Damasceno¹, Fabio Montanha Ramos¹, José Barbosa Dias Jr², José Remo Ferreira Brega²

FESURV – Universidade de Rio Verde
UNIVEM – Centro Universitário Eurípides de Marília

eduardodamasceno@uol.com.br ; remo,@fundanet.br; [@yahoo.com.br](mailto:juniorddos,fa_montanha)

Abstract

With the increasing innovation of the multimodal technology for interface of virtual environments the necessity to have a light computing environment to accept and not loss de interaction created.

To create a way to including a speech resource in the systems without a misinteraction is the propose of this work. With this premise the work is to elucidate some practical and techniques of programming for virtual environments making the comparison between two libraries of three-dimensional environment generation: Java 3D and the World Toolkit, and as these libraries holds meetings of the libraries Java Speech and Microsoft Speech.

1. Introduction

The information technology advances the platforms already dreamed by the man, mainly the levels of accessibility of the information. The interfaces man-computer is each time more complex and robust, needing an onslaught in new technologies access.

The new generation of interfaces man-computer will be an interface of easy learning and high accessibility detaching the use of the voice interfaces, of which would be the equivalent to third part of all the interfaces created for computational environments of high complexity. This is for two main reasons: at first the easiness in the learning of the environment through the use of in agreement voice [7] and second for the freedom that the user can have, as [5] that it describes that the generation and the use of messages or commands said and interpreted for the computer, or either, synthesis and recognition of the voice human being, allowing that the eyes or hands are free for the accomplishment of other tasks.

2. Approach

However, exactly with access facilitated to the set of information the user will be able to feel an anxiety or frustration in the use of the interfaces that if use resources of 3D images or complex information, as it detaches [4].

The felt of perception of the human being is the most complex and complete system of perception of the animals; in it the being can identify variables of the

environment and relate the audible and visual information in intention to understand what it is transferred to around away [6].

With the use of the technology of speak services favoring the usability of the system, the resources placed for the processing, consults the information, and visualization of the information for the user mechanism of recognition and synthesis of speech finish if attenuating in function of the allocation it. From there it has broken the premise of this work that considers the study of the programming techniques and of computational environment development that they use the services of speak without fall down of performance, or with a tolerant fall for the user.

Eyes and brain work together to perceive and to interpret the information of the world making with that the user feels itself entertained with the environment, but when one of them has a delay, or if advances one of the other, brain perceives the misalignment and wakes up the user taking off it of the immersed able and it makes it to attempt against for the problem of synchronism between the appearance and the audio one.

To study notice the behavior of environments created in two languages Java and C/C++ using the resource of speak through the use of the libraries Java Speech and Microsoft Speech inside of a virtual environment to define a set of issues for future developments of this technology in environments of virtual reality.

For in such a way the mechanism of synthesis and recognition of it speaks used was IBM compared ViaVoice and with the mechanism of the Microsoft Speech. It was developed the 3D environments with the application using libraries of generation of virtual environment: Java 3D (SUN) and WorldToolKit (Sense 8).

In this work it was opted in using a mechanism of speech recognition that facilitated the implementation of the service in the virtual environment, and in accordance with the described issues for [3], the option of choice of IBM ViaVoice was the one that more if it proved applicable, for being a steady system and to possess a set of libraries of programming for Java and C/C++.

As same way to prove the study, uses a Microsoft Speech Library to develop another environment with the same characteristics [1].

There are limitations in all libraries of recognition mechanisms will be found detailed in [2], and discuss in the [1].

Previous works are distinguished for the application of the recognition of voice and synthesis using its basic form without if inside worrying about the interactive and immersion of the environment, as it is the case of the described applications in [1].

The process of speech recognition is different of the understanding of speech, the speech understanding of is beyond the speech recognition in which exists the translation of the signal of speaks for a text, while in the understanding of it speaks is generated an action for what he was recognized.

Such mechanism to be able to interact with the virtual system is necessary to use libraries API (Application Programming Interface) that make the work of interface (linking) between the involved parts.

As described by [1], the use of the Java extensions for recognition speech and synthesis speech use the Java Speech API, but this today it is predominant in the development of applications of accessibility to the user in the Java language using the technology VoiceXML and Java Speech.

The Library of programming for Windows is the Microsoft Speech, contends an interface of called programming of Speech API. This is a library that proves the resources of synthesis and based recognition of speech for applicatory in the platform. This set of programming routines consists of the set of archives DLL that it contains the codes of programming necessary to give functionality of the library as a whole.

Any programming language can be used for accessing the libraries of voice of this development group, however it should be taken into account that the more close the language is of the format of development of Microsoft Speech, faster it will be the final application.

This work inherent terms are elucidated the technique of voice services, and between the techniques it can stand out that the way of implementing of the speech services (so much the speech recognition, as the speech synthesis) it depends directly on the attribution that the system should give to the user.

3. Discussions

For the creation of a virtual atmosphere totally immersed it is necessary the inclusion of resources interaction multimodal and one in the ways most accessible of non-conventional interaction it is the use of the recognition technology and speech synthesis.

For the development of applications where is necessary this resources it should be taken some precautions in the choice of the development platform, because same choosing a platform in detriment of other, it can be observed that will always exist like this use

restrictions that they can be noticed by the user, decreasing the immersion of the same.

In the ambit of choosing among platform this work tries to demonstrate some until then characteristics obscure of two development languages. The computer language Java and C/C++.

When observing that the language Java if better interaction with its platform of development of three-dimensional environments in Java 3D.

When it is developed in ambient Java, there is not the concern need with references or messages, not even with its, because Garbage Collection's properties do with that the virtual machine makes the whole arduous work.

When its working directly with the programming with MS-Windows through MVC++ is necessary to understand as the operating system he/she works simulating orientation technology to objects through the changes of messages.

It can be observed that the processing of the virtual environment in Java 3D is faster than the processing in World Toolkit, where it is analyzed the time of rendering pictures, where it is also observed that Java loses less resources when allocating a virtual object, in the in case it tests a change humanoid in different positions of the virtual space.

As the subject of the division of the execution flow once again the language Java prints an advantage on the language C/C++, because it obtains about 7% more of agility in the processing.

It is possible to obtain larger processing when the flow is it divides between the generation of the virtual environment and the recognition mechanism; this so much in Java as in MVC++, but the gain is not so worthless.

The fact that more it denotes the relevance of the research it is that is not necessary to acquire a mechanism of generation of virtual environment like World Toolkit of Sense 8 to generate a three-dimensional environments because the resources of Java 3D can be more worthless.

And when it is necessary the inclusion of recognition resources and speech synthesis the loss of acting of Java it is more constant, what can mean that a virtual environments with speech recognition developed in Java with the technique of correct programming tends to have the low delays this compared to the same environment developed in MVC++ with World Toolkit.

4. References

- [1] DAMASCENO, E.F. & BREGA, J. R. F.; "*Implementação um ambientes virtual conduzido por voz em Java*", IV.SBGames – Simpósio Brasileiro de Jogos de Computador, Curitiba, PR 2004.
- [2] SUN. Java™ Speech API Programmer's GuideVersion 1.0. 1998, <www.java.com>

- [3] PIZZOLATO, E. B. e REZENDE M.N. “*Issues to Consider when Adopting Commercial Speech Interface in Virtual Worlds*”, in Proceedings of Symposium on Virtual Reality, Ribeirão Preto, SP 2003.
- [4] PRESSMAN, R.S. “*Software Engineering: A Practitioner’s Approach*”, 5^a ed. McGraw-Hill, New York, 2002.
- [5] TATHAM, M.. “*Data Retrieval Using Voice*” 1995. <<http://www.essex.ac.uk/speech/pubs/presents/iee-96/iee-96a.html>> 15 Jul. 2003.
- [6] TORTORA, G, “*Structure and interpretation of visual languages*”. New York, 1990..
- [7] ZELTER D. & JOHNSON M. B. “*Interacting with Virtual Environments*”, Ed. John Wiley & Sons, New York, 1994.