CoRSAIRe – Combination of Sensori-motor Rendering for the Immersive Analysis of Results

Brian FG Katz[1], Olivier Warusfel[2], Patrick Bourdot[1], and Jean-Marc Vezien[1]

[1] LIMSI-CNRS
BP 133
91403 Orsay, France
brian.katz@limsi.fr
{first}.{lastname}@limsi.fr

[2] IRCAM
1, place Igor-Stravinsky
75004 Paris, France
Olivier.Warusfel@ircam.fr

ABSTRACT

The CoRSAIRe project (« *Combinaisons de Rendus Sensori-moteurs pour l'Analyse Immersive de Résultats* », or Combination of sensori-motor rendering for the immersive analysis of results) aims at significantly enhancing currently existing interfaces in scientific applications by introducing multiple sensori-motor channels, so that the user will be able to see, hear, and touch the data itself. The project focuses on two well-defined application areas: Fluid Mechanics and Bio-informatics. Stemming from an in-depth confrontation between current observation methods in these fields and existing Virtual Environment interaction techniques, new interaction concepts, paradigms, and concrete solutions are being designed and tested on real-world cases. Such an effort is inherently interdisciplinary and gathers experts from Virtual Reality with knowledge in distributed computing, graphics, audio, and haptic, together with application specialists and experts in ergonomics.

1. INTRODUCTION

The goal of the CoRSAIRe project is to develop new ways of interacting with large or complex digital worlds. The project aims at significantly enhancing currently existing interfaces by introducing multiple sensori-motor channels, so that the user will be able to see, hear, and touch the data itself (or objects derived from the data), thus redefining conventional interaction mechanisms. Such a research effort involves a paradigm shift, because many well-established visualization-oriented software packages exist to analyze the large spectrum of available data types: thus, creating a completely innovative sensori-motor interface would seem a daunting task. This project focuses on two well-defined application areas (Fluid Mechanics and Bio-informatics) with which collaborations are put into place with end-user partners.

A major facet of the project regards how the scientist is able to explore, analyze, and understand large complex datasets. The complexity of the representation in the two disciplines mentioned lies on the one hand in the number of correlated variables to analyze simultaneously and on the other hand, in the many parameters the user must control to successfully drive his analysis. In the CoRSAIRe project, one identified goal is to allow the user to interact with the virtual data in real-time on the evolution of the studied phenomena (to correct, target, modify, annotate...), according to information that he/she perceives. The multimodal management of the interaction is based on the exploitation of stereoscopic visualization, three-dimensional audio, and haptic feedbacks. When considering multimodality, an important and non-trivial issue is the relation between the rules and principles allowing the decision of how the information should be distributed on the different modality channels and how user commands are made available on the different interaction modalities.

On the other hand, in contrast to the visual or haptic modalities, the sonification of data brings a more global comprehension of the information through full 3D reproduction, while also providing an improved representation of the aspects of temporal dependence. Through this method of information presentation, the human capacity for auditory analysis, able to extract the periodicities or the sophisticated constructions over long durations, is particularly well adapted. Development of the sonic representation of large data structures involves several key points of research and investigation. The current methodologies for interactive sonification in the two domain applications are very different. The basis for each shall be discussed in the following sections.

2. MULTI-MODAL DISTRIBUTION

The implementation of multimodal rendering is considered in two parts: the analysis of the end-users' needs and the conception of appropriate multimodal interactions; and the design of hardware and software technologies allowing these interactions.

Rather than attempt designing a fully independent multimodal distribution engine [3], the technological approach used is centered on a "multimodal supervisor." Interactions are designed by the developer of the application, who chooses the paradigms for all the tasks. When the user wants to perform a specific task, the application proposes an interaction to the supervisor which is then in charge of validating, modifying, or specifying it. This choice depends on: a set of distribution rules, the context of the task, the state of the current rendering and the command of the user. The criteria that could influence the rendering must be known at all time. The dynamic context is observed by a module that translates various state variables (tracking data, etc.) into meaningful parameters: where the user is and what he is doing, how the scene is organized, etc. Another module is in charge of observing the static and dynamic rendering capacities of the application (available modalities in the application and charge of these modalities). Regarding user commands, when he has decided to do a task with a specific paradigm, the supervisor must take into account this high level command: which means it should validate the interaction, even if it is not the best choice, unless it is impossible or dangerous.

3. APPLICATIONS

The two main scientific application fields of this project, which are Computational Fluid Dynamics and Molecular Bio-informatics, are directly concerned with the crucial problem of generation and processing of large volumes of complex data. These data are generally the results of experiments, simulations, or numerical predictions. For example, in Bio-informatics, the scenes (DNA molecules) can be made up of several million atoms whose interactions are dynamically controlled by the research operator. In Fluid Mechanics, classical examples study the characteristics of a non stationary three-dimensional flow for which a multitude of parameters are calculated at every point. In general, geometrical structures consisting of millions of primitives are essential for a good visual spatial resolution of an instantaneous field. This corresponds to a volume of data of about 10 to 100 Megabytes per scalar field for a relatively small scale problem. The study of the temporal evolution of non stationary fields requires the production of tens, if not hundreds, of such datasets. This leads to gigabytes of unstructured data.

3.1. Fluid Mechanics

Physicists, on their quest to find characteristics of non-stationary 3D experiments, often perform computed simulation of their real experiments. These simulations output massive grids of numerous computed vector or scalar fields (pressure field, velocity, and other derived fields such as vorticity and helicity) for each time step of the simulation. Experts used to (and still often) study these simulation results in the form of cutting planes (i.e. 2D slices of the whole data volume) which are comparable to experimental data, or 3D static representations such as isosurfaces (see Figure 4). An objective of CoRSAIRe is to find news ways of virtually exploring massive grids coming from Computational Fluid Dynamics (CFD) simulations, to enhance the understanding of the studied phenomena.

CFD calculations produce massive amounts of temporary data as they calculate the temporal evolution of the fluid system. It is typical for the final results to be only a sub-sample of the calculation. This results in a time varying resolution of the data. For example, a simulation is run with data calculated at time steps of 0.0025 sec (or 400 Hz). The data at several individual locations of interest in the volume are stored at full time resolution, a sampling rate of 400 Hz. A collection of data planes are stored at a frequency of 60 Hz, while the entire data volume is stored at a data sampling frequency of 30 Hz. While the last two datasets are well suited to visual and haptic exploration, the first dataset is not readily translatable to this representation in an immersive environment. In contrast, the singular point time series data is well-suited to an audio representation. Preliminary studies have been carried out using a process of up-sampling and pitch expanding, using a phase vocoder, to bring the low sample rate numerical data into the full audio spectra. Various degrees of pitch expansion result in different characteristics of the data becoming more or less perceptible. The auditory presentation of these time series data, spatially located within the immersive environment and correctly located with respect to the static isosurface visual representation, provides an incredibly new and interesting perspective on the data set. This metaphor presents a continuous sonification loop, where sound sources can be activated depending on user controls, proximity, etc.

Additional and more abstract data sonification metaphors are being explored, which are based on the 2D and 3D static representations of the data field. Such abstract metaphors could be envisioned as continuous sonification of selected aspects or dimensions of the data field. An audio spotlight could be employed to spatially select data field zones.

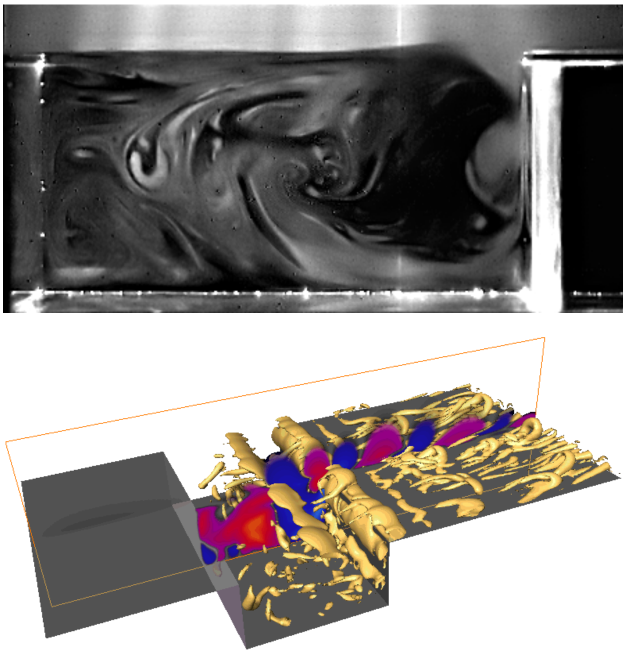

Figure 4. Fluid mechanic e*xperimental data snapshot (Particle Image Velocity, PIV) for 2D analysis of dynamic processes (above) and numerical simulation of a similar condition showing isosurface and cutting plane (below).*

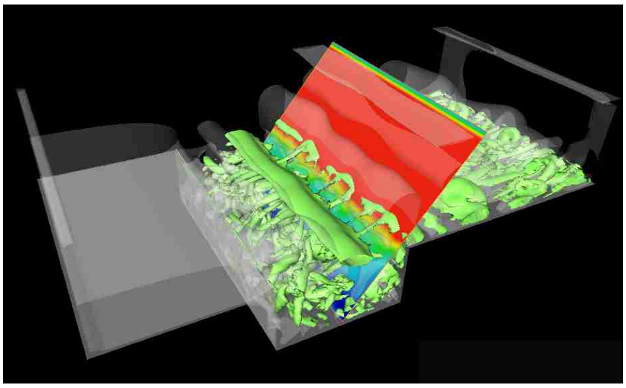

Figure 5. *Numerical simulation of experimental configuration showing isosurface (green), cutting plane, and haptic guide (transparent isosurface).*

A typical VR exploration scenario (Figure 5) would be to visualize multiple relevant isosurfaces of a given dataset, while being haptically guided by a significant and complex invisible isosurface. A 3D cutting plane is attached to the cursor, showing additional information on other fields during the exploration. Finally, a set of points, representing specific ROI where temporal series were simulated and recorded, are spatially auralized with 3D sound, to provide information on the temporal evolution of the static field observed.

3.2. Bio-informatics

The main objective of the Bio-informatics application is to propose, implement, and evaluate a new approach for the protein-protein docking problem (prediction of protein-protein association) based on the precept that such complex simulations could be more effective if user-driven and the biologist is considered in the docking process loop. Human perceptual skills are a valuable asset for 3D-pattern analysis, recognition, and mining tasks. This is also true in the decision-making phase providing in consequence an increase in performance, both in time and in the quality of the docking computation. Advanced virtual reality multimodal interfaces combining gestures, speech recognition, audio synthesis and haptics enable the biologist to act naturally in the simulation process and to skim through the set of possibilities and make a fast selection of biologically interesting protein complexes. User interaction with the simulation is also carefully recorded in the form of annotations (including vocal), so that all the useful information relative to the history of the docking procedure is recorded.

The majority of current biological research focuses on the relationship between structures and functions. But, because the majority of proteins provide their function in the form of complexes, it is the structure of the complex which is actually the object of interest. Various works have been carried out concerning immersive docking: STALK [1], ARCdocking [9], etc. Nevertheless the obtained results are still too specific or insufficient. These studies consider the user more as an observer rather than an actor. The approach presented here is more human-centered and addresses the design of user-oriented multimodal immersive docking systems (Figure 1).

The computational choice of the molecular representation model is crucial for the quality of the docking. Indeed, it is based, first of all, on the surface complementarity of the implied proteins. Existing representations, e.g. "Van der Waals" one, or the "wire" model, are both unsuitable for the docking problem, because they contain too much inner information and not enough surface information. Surface-based models [8] enable better access to the surface of the protein, being more adequate for the search of binding sites between two proteins. But, these models are too complex to use in a virtual reality application. Therefore a visual simplification of the molecular representation is needed. Initial investigations [7] exploit a Voronoï model adapted to the protein-protein docking problem.

Aside from the selected machine representation of molecules, one must also carefully consider the visual, audio, and haptic representations that will provide maximum benefit during interaction. These are closely related to the nature of such interaction, and should provide meaningful feedback on the data involved in docking computations.

In order to solve problems of "key-lock" type, haptic feedback can be as useful as visual feedback [11]. Previous research has validated haptic feedback in docking applications, for example where the force feedback system is used to simulate the interaction between a ligand, considered as a probe of force and torsion, and a receiver [10]. The results obtained show that the chemists found the position of minimal potential energy in half the time relative to systems only based on visualization. Nagata [5] proposed a similar system for the ligand-protein docking with the ligand modeled as a potential probe. The problem is that the test probe was composed only of one sphere with only one charge, whereas a ligand must contain at least ten atoms having each one its own charge and a protein must contain at least several hundred atoms having too each one its own charge.

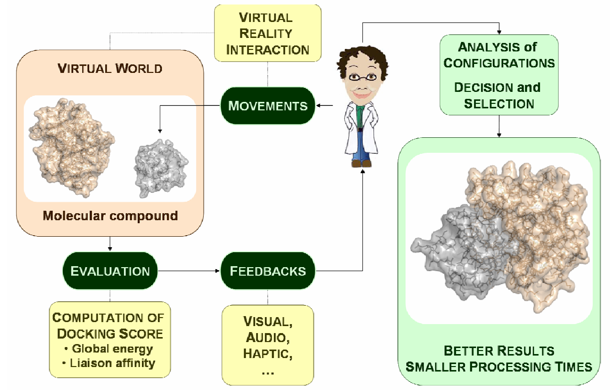

Figure 1. *Immersive docking prototype principle.*

The protein docking database includes the ensemble of compatibility interactions between each surface element of one protein (using the Voronoi model) to each and every surface element of the second. These interactions are of numerous types, and can be translated to force vectors proportional to the distance between surface elements. One proposed auditory representation of this data is to consider the sum of "forces" from the probe protein on each element of the receiver protein (the choice being relatively arbitrary). This is an extension of a simpler sonification metaphor for vector fields [2]. The resulting collection of force vectors have a magnitude, direction, and can also be classified as attractive or repulsive with regards to their docking potential (see Figure 2). Such data can be 3D spatialized to help the user to localize regions of interest (ROI) where, for example, docking is highly problematic.

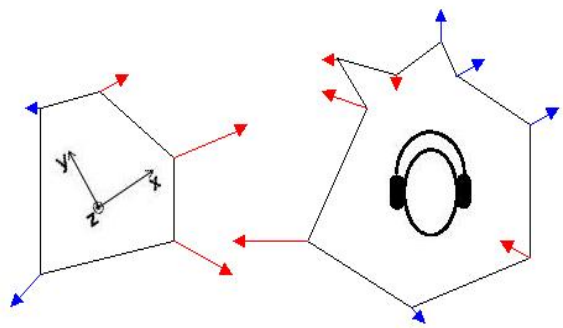

Figure 2. *Schematic of force sonification concept, showing the user's geometrical location.*

User interaction is continuous in nature, and results in a continuous evolution of the data field. Therefore, a continuous data sonification metaphor is needed. Each element can be considered as an audio generator, where the frequency and level of the generator are controlled by the magnitude (thereby eliminating the presence of negligible interactions and emphasizing the strongest) and the nature of the signal is dependant on the direction or attractive/repulsive character of the result (whereby conflicting arrangements are "harsh" and compatible arrangements are the opposite). The spatial distribution of the sonified data is relative to the geometrical structure of the receiver and a user centered point of reference (see Figure 3).

An alternative sonification metaphor consists in generating an aural representation of more abstract data, such as global averages of various interaction compatibilities. This approach requires the development of more sophisticated mapping schemes to convert the data to audible information [4][6][12]. Each resulting sonified data stream can be spatially separated in order to optimize intelligibility of the data stream.

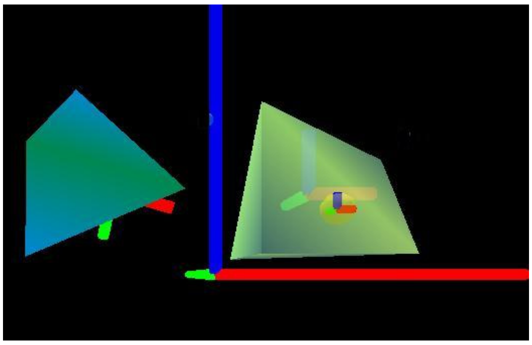

Figure 3. *Sonification model mock-up showing 2 "protein" polyhedrons. The user is located within the fixed object, while the target is manipulated around the user.*

4. CONCLUSIONS

The CoRSAIRe project has many important research challenges, from technological problems of multi-sensory integration and management of massive data, to design issues of new paradigms of VR interaction which can significantly aid the scientist end-users of the targeted applications. With such applications, a major problem is how the scientist is able to explore, analyze, and understand these datasets. The complexity of the representation in the two mentioned disciplines lies in the number of correlated variables to analyze simultaneously and in the many parameters the user must control to successfully drive his analysis. One identified goal of CoRSAIRe is to allow the user to interact with the virtual data in real-time (to correct, target, modify, annotate, etc.), according to information that he perceives. The multimodal management of the interaction is based on the exploitation of stereoscopic visualization, haptic, and three-dimensional audio feedbacks. It is a key assertion of CoRSAIRe that an efficient multimodal interaction will prove highly beneficial to the users of immersive tools dedicated to scientific analysis.

Several metaphors have been proposed and are in the course of development for the sonification of abstract, continuous, not-structured data fields. Emphasis is on continuous sonification, rather than alert messaging, as the goal of the sonification process is to increase the understanding of the data field by the application's users through continuous interaction in an immersive virtual environment. The design of sonification building blocks, which are data independent, will allow for a flexible system, where metaphors can easily be explored, selected either through user commands or via a multimodal supervisor, with knowledge of the underlying data structure and the strength of each metaphor.

5. ACKNOWLEDGEMENTS

CoRSAIRe is a 3 year research project supported by ANR (French National Research Agency) labeled in December 2005. Lead by LIMSI-CNRS (VR, Multimodal interaction, 3D audio, and Fluid Mechanics aspects), the CoRSAIRe consortium is also composed of three other research teams: IRCAM-CNRS (VR and 3D audio expertise), EBGM - INSERM (Bio-informatics application), ECI - Univ. Paris 5 (analysis of user needs and Ergonomics studies on VR applications), and HAPTION S.A (Haptic interfaces).

6. REFERENCES

[1] D. Levine, M. Facello, P. Hallstrom, G. Reeder, B. Walenz, and F. Stevens, "STALK: An Interactive Virtual Molecular Docking System," *IEEE Comput. Sci. Eng*., 4, 55-65, 1996.

[2] E. Klein and O.G. Staadt, "Sonification of Three-Dimensional Vector Fields," in *Proceedings of the SCS High Performance Computing Symposium*, 2004.

[3] G. Bouyer, P. Bourdot, and J.-M. Vézien, "TOTEM: An Engine for Task-Oriented Multimodal Rendering in Virtual Environments." in *Proceedings of the HCI International Conference (HCII)*, Vol. 9: Advances in Virtual Environments Technology: Musings on Design, Evaluation, & Applications. Las Vegas, USA, July 2005.

[4] G. Kramer, "Some organising principles for representing data with sound." *Auditory display: Sonification, audification, and auditory interfaces (Proceedings Volume 18, Santa Fe Institute Studies in the Sci),* G. Kramer (Ed.), (pp. 79-94). Reading, MA : Addison-Wesley. 1994.

[5] H. Nagata, H. Mizushima and H. Tanaka, "Concept and prototype of protein-ligand docking simulator with force feedback technology," *Bioinformatics,* Vol.18 No.1 p.140-146, 2002.

[6] J. Anderson, "Creating an Empirical Framework for Sonification Design," in *Proceedings of ICAD 05*, Limerick, Ireland, July 6-9, 2005.

[7] J. Bernauer, *Utilisation de la tesseletion de Voronoï pour l'étude des complexes protéine-protéine*, Thesis, Institut de Biochimie et de Biophysique Moléculaire et Cellulaire, Orsay France, 2006.

[8] M.L. Connolly, "Solvent-accessible surfaces of proteins and nucleic acids," *Science*, 19, 2003.

[9] N. Ray, X. Cavin, J.-C. Paul, and B. Maigret, "Intersurf: dynamic interface between proteins," *Journal of Molecular Graphics and Modelling*; Volume 23, Issue 4, January 2005, p.347-354.

[10] R. M. Taylor II, "Scientific Applications of Force Feedback: Molecular Simulation and Microscope Control" *SIGGRAPH'99*.

[11] S. Birmanns and W. Wriggers, "Interactive fitting augmented by force-feedback and virtual reality," *Journal of Structural Biology*, 144, 2003, p.123-131.

[12] T. Hermann and H. Ritter, "Listen to your Data: Model Based Sonification for Data Analysis," in *Advances in intelligent computing and multimedia systems*, M.R.Syed,Ed. 1999.