



UNIVERSITAT POMPEU FABRA

# Computer-Supported Cooperative Work for Music Applications

Author: Álvaro Mendes Barbosa

April 2006



Dipòsit legal: B.48971-2006  
ISBN: 978-84-690-3933-5

Dissertation submitted to the Department of Technology of the Pompeu Fabra University for the Program in Computer Science and Digital Communication, in partial fulfillment of the requirements of the degree:

Doctor per la Universitat Pompeu Fabra with Mention of European Doctor

Dissertation directed by Dr. Xavier Serra and co-directed by Dr. Sergi Jordà

Universitat Pompeu Fabra  
Departamento de Tecnologia  
Estació de França  
Passeig de Circumvallació, 8  
08003 Barcelona, España



Research leading to this Dissertation was conducted by the author at:



This Doctorate Research Work was supported through the award of a Doctorate Scholarship by:

**FCT** Fundação para a Ciência e a Tecnologia

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR Portugal

(SFRH/BD/5192/2001)

The Author is affiliated with:





To Sofia





# Abstract

This dissertation derives from research on musical practices mediated by computer networks conducted from 2001 to 2005 in the Music Technology Group of the Pompeu Fabra University in Barcelona, Spain. It departs from work carried out over the last decades in the field of Computer-Supported Cooperative Work (CSCW), which provides us with collaborative communication mechanisms that can be regarded from a music perspective in diverse scenarios: Composition, Performance, Improvisation or Education.

The first contribution originated from this research work is an extensive survey and systematic classification of Computer-Supported Cooperative Work for Music Applications. This survey led to the identification of innovative approaches, models and applications, with special emphasis on the shared nature of geographically displaced communication over the Internet. The notion of a Shared Sonic Environments was introduced and implemented in a proof-of-concept application entitled Public Sound Objects (PSOs).

A second major contribution of this dissertation concerns methods that reduce the disrupting effect of network latency in musical communication over long distance networks. From laboratorial experimentation and evaluation, the techniques of Network Latency Adaptive Tempo and Individual Delayed Feed-Back were proposed and implemented in the PSOs prototype.

Over the course of the PSOs development other relevant and inspirational issues were addressed, such as, behavioral-driven interface design applied to interface decoupled applications, the overcome of network technology security features and system scalability for various applications in audio web services.

Throughout this dissertation conceptual perspectives of related issues to computer-mediated musical practices dissertation were widely discussed, conveying different standpoints ranging from a Psycho-Social study of collaborative music processes to the Computer Science and Music Technology point of view.

# Resum

Aquesta tesi recull la recerca al voltant de les pràctiques musicals mitjançant xarxes d'ordinadors realitzada al Grup de Tecnologia Musical de la Universitat Pompeu Fabra a Barcelona entre l'any 2001 i el 2005. Parteix del treball dut a terme durant la última dècada dins del camp del Treball Cooperatiu amb Ordinadors (Computer-Supported Cooperative Work, CSCW) el qual aporta els mecanismes de col·laboració els quals, des de un punt de vista musical, poden ser estudiats en diversos escenaris: composició, interpretació, improvisació i educació.

La primera contribució d'aquest treball és un anàlisi exhaustiu i una classificació sistemàtica del Treball Cooperatiu amb Ordinadors per Aplicacions Musicals. Aquest anàlisi es va centrar en la identificació de propostes innovadores, models i aplicacions, amb un especial èmfasi en la natura compartida de la comunicació mitjançant internet. El concepte d'Entorns Sonors Compartits va ser presentat i implementat en una aplicació prototip anomenada Public Sound Objects (PSOs).

La segona gran contribució d'aquesta tesi consisteix en l'estudi del possibles mètodes per reduir les interrupcions degudes als retards inherents en la comunicació musical entre xarxes molt allunyades. A partir de l'experimentació i avaluació al laboratori les tècniques Network Latency Adaptive Tempo i Individual Delayed Feed-Back van ser definides i implementades dins del prototip PSOs.

Al llarg del desenvolupament del PSOs es van haver de resoldre altres problemes, com per exemple, el disseny d'interfícies en funció del comportament per a aplicacions amb interfícies desacoblades, la superació dels diversos sistemes de seguretat de les xarxes informàtiques i les possibilitats d'escalabilitat de diverses aplicacions d'àudio per a web.

Durant l'elaboració d'aquesta tesi es van discutir diferents perspectives per resoldre problemes relacionats amb la pràctica musical mitjançant ordinadors, aplicant diferents punts de vista provinents de l'estudi psicosocial dels processos de col·laboració musical al món de la informàtica i de la tecnologia musical.

# Acknowledgments

I would first like express my gratitude to Professor Francisco Carvalho Guerra for his trust and encouragement over the last five years, as well as the opportunities he has provided for me and all my colleagues at the School of Arts at the Portuguese Catholic University in Porto, Portugal.

I am deeply grateful to my doctorate advisors Xavier Serra and Sergi Jordà for their guidance, unconditional support and most of all for providing a role-model of character and knowledge, which will always be a reference to me.

For their support, fruitful discussions about my ideas and collaboration in the work which gave birth to this dissertation, I am also particularly grateful to:

Josep Blat, Alexandro Ramirez, Martin Kaltenbrunner, Günter Geiger, Alexander Carôt, Fabien Gouyon, Pedro Cano, Diego Dall'Osto, Perfecto Herrera, Ross Bencina, Marcos Alonso, Rafael Ramirez (Pompeu Fabra University, Barcelona);

Jorge Cardoso, João Seabra, Paulo Ferreira-Lopes, Guilhermina Castro, Joana Cunha e Costa, Daniela Coimbra, Luís Gustavo Martins, Carlos Barreiros, Mafalda Barbosa, Mariana Madaíl and Kurt Stewart (Portuguese Catholic University, Porto).

For extremely helpful discussions, advices and inspiring ideas related with the main topics of my dissertation, I am tremendously indebted to my friend Carlos Baquero Moreno (Minho University, Portugal).

For their acknowledgement and interest in my work I would like to express my gratitude to Curtis Roads (University of Santa Barbara), Antonio Camurri (University of Genoa), Chris Chafe (Stanford University), Atau Tanaka (Sony CSL Paris), Dante Tanzi (University of Milan), Kiyoshi Furukawa (Tokyo National University), Jason Freeman (Georgia Institute of Technology), Scot Gresham-Lancaster (Cogswell College Sunnyvale, CA), Suguru Goto (Paris VIII University), Nicolas Collins (Editor of Leonardo Music Journal) and Leigh Landy (Editor of Organised Sound).

For their help, proof-reading and bearing with me, while writing my thesis, I would like to thank my friends Carla Almeida, Cristina Sá, Sahra Kunz, Helena Figueiredo, Cinthia Ruiz and Joana Martins.

I also owe a great debt of thanks to my colleagues at the Portuguese Catholic University, with whom I've shared a rich work experience over the last years: José Paulo Antunes, Teresa Macedo, Gonçalo Vasconcelos, Maria Lopes Cardoso, Amílcar Sousa, José Miguel Cadilhe, Miguel Lobo, Carlos Caires, Luis Teixeira, Paulo da Rosária, Hélder Dias, Adriano Nazareth, Baltazar Torres, Mónica Monteiro, Miguel Rodrigues and Jaime Neves.

I would like to mention my appreciation to the people that formerly trusted and encouraged me in my professional career: Conego Ferreira dos Santos, Luis Proença, Tiago Azevedo Fernandes and Armando Batista.

I would also like to acknowledge that my Doctorate research was funded by the Portuguese institution Fundação para Ciência e Tecnologia, through the award of a Doctorate Scholarship (SFRH/BD/5192/2001).

# Table of Contents

<b>Abstract.....</b>	<b>i</b>
<b>Abstract (Catalan) – Resum.....</b>	<b>ii</b>
<b>Acknowledgements.....</b>	<b>iii</b>
<b>Table of Contents.....</b>	<b>v</b>
<b>List of Figures.....</b>	<b>x</b>
<b>List of Graphics and Tables.....</b>	<b>xiv</b>
<b>Chapter 1 - Introduction .....</b>	<b>1</b>
<b>1.1 Motivation.....</b>	<b>2</b>
<b>1.2 Objective of this Dissertation .....</b>	<b>3</b>
<b>1.3 Structure of this Dissertation .....</b>	<b>3</b>
<b>Chapter 2 - Survey of Computer-Supported Cooperative Work for     Music Applications .....</b>	<b>5</b>
<b>2.1 Computer-Supported Cooperative Work.....</b>	<b>5</b>
2.1.1 Operation Modes in CSCW.....	7
2.1.2 Synchronous and Asynchronous Modes in CSCW .....	8

2.1.3 The CSCW Classification Space.....	10
2.1.4 Shared Virtual Environments .....	10
<b>2.2 Computer Mediated Communication and Networked Music .....</b>	<b>13</b>
2.2.1 Collaboration in Music from a psycho-social perspective .....	14
2.2.1.1 Towards a New Social Space of Music creation.....	16
2.2.1.2 Csikszentmihalyi’s Creative Person in Shared Soundscapes .....	17
2.2.1.3 Csikszentmihalyi’s Domain of Shared Soundscapes .....	18
2.2.1.4 Csikszentmihalyi’s Field in Shared Soundscapes .....	19
2.2.2 Redefining the Acoustic Community for Music and Sonic Arts.....	20
2.2.3 Networked Music as a Research Topic .....	21
2.2.3.1 Landmarks in Networked Music Research .....	23
<b>2.3 Systematic Study of Networked Music Systems.....</b>	<b>28</b>
2.3.1 Early Experiments with Musical Networks.....	29
2.3.2 Geographical Displacement in Music Communication.....	34
<b>2.4 Networked Music Systems Overview .....</b>	<b>37</b>
2.4.1 Co-Located Musical Networks.....	39
2.4.2 Music Composition Support System.....	40
2.4.2.1 On-line Music Recording Studios .....	41
2.4.2.2 Experimental Collective Composition Systems .....	47
2.4.3 Remote Music Performance Systems.....	53
2.4.3.1 Tele-Presence Systems .....	53
2.4.3.2 Collaborative Performance Systems .....	56

2.4.4 Shared Sonic Environments .....	65
2.4.4.1 On-Line Improvisation.....	66
2.4.4.2 The Time-Scales of a Permanent Event .....	67
2.4.4.3 System Implementations .....	68
<b>2.5 Chapter Conclusions.....</b>	<b>71</b>
<b>Chapter 3 - Networked Music Practice Topologies .....</b>	<b>73</b>
<b>3.1 Networked Models for Collaborative Music Practice .....</b>	<b>73</b>
3.1.1 Common Network Protocols, Architectures and Models.....	74
3.1.1.1 Reliability and Quality of Service .....	76
3.1.1.2 Network Communication Models .....	77
3.1.1.3 Decentralized Communication Environments.....	80
3.1.2 General-Purpose Models for Music Collaboration.....	82
<b>3.2 Towards an Ubiquitous Virtual Music Instruments .....</b>	<b>87</b>
3.2.1 Traditional Music Instruments VS Virtual Music Instruments .....	87
3.2.2 Nomadic Music Instrument Model.....	91
<b>3.3 Multimodality and Networked Music .....</b>	<b>92</b>
3.3.1 Personal Digital Assistants (PDAs) as Music Controllers .....	95
3.3.2 The ReactTable* .....	97
<b>3.4 Chapter Conclusions.....</b>	<b>101</b>
<b>Chapter 4 - Internet Acoustic Communication Facets .....</b>	<b>103</b>
<b>4.1 The Perception of Internet Acoustics .....</b>	<b>104</b>

4.1.1 Latency Tolerance in Music Performance.....	105
4.1.2 Latency Adaptive Tempo and Dynamics .....	106
4.1.3 Individual Delayed Feed-Back.....	114
<b>4.2 Web services and Acoustic Applications.....</b>	<b>119</b>
4.2.1 Semiautomatic Ambiance Generation On-Line .....	120
4.2.2 Data Sonification on Demand .....	122
<b>4.3 Chapter Conclusions.....</b>	<b>125</b>
<b>Chapter 5 - The Public Sound Objects: A System Prototype for Experimental Research.....</b>	<b>127</b>
<b>5.1 Community Music and Sound Objects .....</b>	<b>129</b>
<b>5.2 System Overview and Architecture.....</b>	<b>131</b>
5.2.1 Web Server.....	133
5.2.2 Communication Layer.....	133
5.2.3 Synthesis and Transformation Engine.....	135
<b>5.3 User Interface .....</b>	<b>137</b>
5.3.1 Multi-Platform Implementation .....	142
5.3.2 Installation Site.....	145
<b>5.4 Distinctive Software Features .....</b>	<b>147</b>
<b>5.5 System Evaluation.....</b>	<b>152</b>
<b>5.6 Chapter Conclusion .....</b>	<b>158</b>



<b>Chapter 6 - Conclusions and Future Work .....</b>	<b>160</b>
--	------------

<b>6.1 Summary of Contributions .....</b>	<b>160</b>
---	------------

<b>6.2 Future Directions .....</b>	<b>163</b>
------------------------------------	------------

<b>Bibliography.....</b>	<b>170</b>
--------------------------	------------

<b>Glossary .....</b>	<b>179</b>
-----------------------	------------

## **APPENDIX A: Published Work by the Author**

Papers in Peer-Reviewed Journals .....	183
Papers in Peer-Reviewed Conferences .....	185
Other Related Publications .....	190

## **APPENDIX B: Companion DVD Video**

Documental Video Essay about the Public Sound Objects System (00:05:13)

# List of Figures

<b>Figure 1.</b> Rodden’s Classification Space for CSCW Applications .....	<b>10</b>
<b>Figure 2.</b> Diagram of the Computer Music interdisciplinarity field proposed by (Moore, 1990). .....	<b>22</b>
<b>Figure 3.</b> Participants of the ANET Summit; Organizers: Chris Chafe (1st left), Jeremy Cooperstock (4th Right), Theresa Leonard (6th Right), Bob Moses (3rd Right), Wieslaw Woszczyk (3rd left).....	<b>25</b>
<b>Figure 4.</b> The League of Automatic Music Composers (Perkis, Horton, and Bischoff, left to right) performing at Ft. Mason, San Francisco 1981. Photo: Peter Abramowitsch. ...	<b>32</b>
<b>Figure 5.</b> Flyer, designed by Rich Gold, from 1979 announcing a regular series of concerts, and showing different network connection topologies between the League computers....	<b>33</b>
<b>Figure 6.</b> A Classification Space Networked Music Systems .....	<b>38</b>
<b>Figure 7.</b> Co-Located Musical Networks in Networked Music Classification Space .....	<b>39</b>
<b>Figure 8.</b> Music Composition Support Systems in Networked Music Classification Space.....	<b>40</b>
<b>Figure 9.</b> Draft Model of Endogenous and Exogenous Creative Trajectory .....	<b>42</b>
<b>Figure 10.</b> Screen Shots of the ResRocket Software showing a structured list of ongoing sessions and a multi-track project view with individual tracks recorded by different users .....	<b>43</b>
<b>Figure 11.</b> Rocket Power Audio system Topology .....	<b>43</b>
<b>Figure 12.</b> TONOS-TC8 Interface.....	<b>44</b>
<b>Figure 13.</b> DigitalMusician.Net User Interface .....	<b>45</b>
<b>Figure 14.</b> eJaming Stage interface .....	<b>46</b>
<b>Figure 15.</b> The VST-Tunnel Plugging Interface.....	<b>47</b>
<b>Figure 16.</b> The FMOL Bamboo Interface.....	<b>48</b>
<b>Figure 17.</b> Screen Shots of the FMOL Software showing the web based tree-structured data based with multiple generation pieces composed by different users .....	<b>49</b>
<b>Figure 18.</b> The Free Sound Project Remix! Tree Interface .....	<b>52</b>
<b>Figure 19.</b> Remote Music Performance Systems in the Networked Music Classification Space .....	<b>53</b>
<b>Figure 20.</b> Diagram for a remote collaborative musical performance for a pianist (using Yamaha Disklavier Pianos) and live electronics performed on a Laptop. ....	<b>56</b>
<b>Figure 21.</b> Diagrams from the TransMIDI System showing possible group topologies .....	<b>58</b>
<b>Figure 22.</b> Musicians at McGill University (Dan Levitin – sax and Ives Levesque – trombone) Jamming with remote Musicians at Stanford University projected on Screen (Alexander Carôt – Bass and Estabin Wilson – sax). .....	<b>60</b>

<b>Figure 23.</b> Screen-Shot from the Peersynth network synthesizer.....	<b>62</b>
<b>Figure 24.</b> Screen-Shot from the Quintet.net client interface.....	<b>63</b>
<b>Figure 25.</b> NINJAM client interface.....	<b>64</b>
<b>Figure 26.</b> Shared Sonic Environments in the Networked Music Classification Space.....	<b>65</b>
<b>Figure 27.</b> Screenshot from the demonstration video Documentary on the Auracle.....	<b>70</b>
<b>Figure 28.</b> Graphical Representation of different computational instances in the Co-Audicle..	<b>71</b>
<b>Figure 29.</b> Transmission Control Protocol (TCP) and the User Datagram Protocol (UDP) .....	<b>75</b>
<b>Figure 30.</b> Centralized Network Model.....	<b>78</b>
<b>Figure 31.</b> Centralized Network Model with Multiple Communicating Servers .....	<b>78</b>
<b>Figure 32.</b> Distributed Network Models.....	<b>79</b>
<b>Figure 33.</b> Group Abstraction for Broadcast Transmission.....	<b>80</b>
<b>Figure 34.</b> Centralized Shared Sonic Environment Model with Local Feed-Back .....	<b>84</b>
<b>Figure 35.</b> Centralized Sonic Shared Environment Model with Local Feed-Back .....	<b>85</b>
<b>Figure 36.</b> Traditional Music Instrument Interaction Model.....	<b>88</b>
<b>Figure 37.</b> Violin Interaction Model.....	<b>89</b>
<b>Figure 38.</b> Virtual Music Instrument Interaction Model .....	<b>90</b>
<b>Figure 39.</b> Nomadic Virtual Music Instrument Model.....	<b>92</b>
<b>Figure 40.</b> Oskar Fischinger’s device for producing light effects. ....	<b>94</b>
<b>Figure 41.</b> Multi-Touch Sensing through Frustrated Total Internal Reflection.....	<b>95</b>
<b>Figure 42.</b> A Compaq IPAQ Running a PD Patch, Gunter Geiger performing with a PDA.....	<b>96</b>
<b>Figure 43.</b> The Public Sound Objects Client Interface running on a PDA.....	<b>97</b>
<b>Figure 44.</b> The ReacTable* architecture (illustration by Ross Bencina).....	<b>98</b>
<b>Figure 45.</b> “TeleSon” Performance September 04, 2005: Chris Brown and Gunter Geiger at ICMC 2005 in Barcelona, Spain (on stage at SGAE auditorium); Martin Kaltenbrunner and Marcos Alonso at Ars Electronica Festival in Linz, Austria (on screen).....	<b>99</b>
<b>Figure 46.</b> Schematics for two Networked ReacTables* (illustration by Ross Bencina).....	<b>100</b>
<b>Figure 47.</b> OSCGroups Communication Model by Ross Bencina .....	<b>100</b>
<b>Figure 48.</b> Ideal Communication Scenario between two globally displaced Cities. ....	<b>106</b>
<b>Figure 49.</b> Experiment on latency tolerance in a simulated studio environment.....	<b>107</b>
<b>Figure 50.</b> Online survey to evaluate the relationship between Musical Tempo and Communication Latency .....	<b>110</b>
<b>Figure 51.</b> Music notation for the song “Sunny” from the Jazz Real-Book.....	<b>111</b>
<b>Figure 52.</b> Normal Feed-Back Topology .....	<b>115</b>
<b>Figure 53.</b> Individual Delayed Feed-Back Topology .....	<b>115</b>
<b>Figure 54.</b> Music notation for the song “Cantaloupe Island” from the Jazz Real-Book .....	<b>116</b>
<b>Figure 55.</b> Online Survey for the evaluation of Individual Delayed Feed-Back Performances	<b>116</b>

<b>Figure 56.</b> Soundjack Interface by Alexander Carôt .....	<b>118</b>
<b>Figure 57.</b> Flow diagram of the Semiautomatic Ambience generation On-Line system. ....	<b>122</b>
<b>Figure 58.</b> Interfaces of the SDM Project for mobile devices .....	<b>124</b>
<b>Figure 59.</b> Architecture of the SDM System .....	<b>125</b>
<b>Figure 60.</b> The Public Sound Objects Architecture.....	<b>132</b>
<b>Figure 61.</b> Entry Screen for PSO Client Version 3 .....	<b>136</b>
<b>Figure 62.</b> Small Fish Visuals by Masaki Fujihata.....	<b>138</b>
<b>Figure 63.</b> First prototype of the PSOs interface developed in Flash.....	<b>139</b>
<b>Figure 64.</b> First Java implementation of the PSOs interface. ....	<b>141</b>
<b>Figure 65.</b> PSOs GUI version released in 2004; Entry screen for nine sound objects and controller with ball tail and real-time network latency measurement. ....	<b>142</b>
<b>Figure 66.</b> Desktop, Touch Screen and PDA Interfaces for PSOs .....	<b>143</b>
<b>Figure 67.</b> PSO Banner embedded at the Home Page for <a href="http://www.abarbosa.org/">http://www.abarbosa.org/</a> .....	<b>144</b>
<b>Figure 68.</b> Mockup design of PSO Installation and the real implementation at Porto School of the Arts, October 2004.....	<b>146</b>
<b>Figure 69.</b> PSOs GUI version released in 2004; Including Multiple Users Graphic representation and other distinctive software features. ....	<b>148</b>
<b>Figure 70.</b> Representation of Impacts VS Triggered Sound without time lag overlap.....	<b>150</b>
<b>Figure 71.</b> Representation of Impacts VS Triggered Sound with time lag overlap.....	<b>151</b>
<b>Figure 72.</b> Representation of Impacts VS Triggered Sound with sound panorama adjustment	<b>152</b>
<b>Figure 73.</b> Locations of the PSO performance on March the 31st of 2005.....	<b>153</b>
<b>Figure 74.</b> João Seabra, Jorge Cardoso and Álvaro Barbosa performing simultaneously with PSOs respectively in Toronto, Porto and Barcelona.....	<b>153</b>
<b>Figure 75.</b> PSOs Installation at NIME 2005 – New Interfaces for Musical Expression Conference, 26-28 of May Vancouver, Canada.....	<b>154</b>
<b>Figure 76.</b> PSOs Installation at ICMC 2005 – International Computer Music Conference, 5-9 of September Barcelona, Spain.....	<b>154</b>
<b>Figure 77.</b> PSOs Trial Installation at Porto School of Arts, 7-14 October 2004 .....	<b>155</b>

# List of Tables and Graphics

<b>Table 1.</b> Maximum delay tolerance for each musician playing at different tempos.....	<b>108</b>
<b>Graphic 1.</b> Self-Test for latency tolerance in individual performance.....	<b>109</b>
<b>Table 2.</b> Evaluation results for Musical Tempo/Communication Latency relationship grouped by musical training levels .....	<b>112</b>
<b>Graphic 2.</b> Evaluation results for Musical Tempo/Communication Latency relationship in the case of Bass/Guitar and Bass/Piano duets.....	<b>112</b>
<b>Graphic 3.</b> Evaluation results for Musical Tempo/Communication Latency relationship in the case of Bass/Percussion duet and a final overall average result .....	<b>113</b>
<b>Table 3.</b> Results from the online Survey on evaluation of Individual Delayed Feed-Back.....	<b>117</b>
<b>Graphic 4.</b> Opinion Pool Characterization.....	<b>155</b>
<b>Graphic 5.</b> Opinion Pool Characterization and question #1.....	<b>156</b>
<b>Graphic 6.</b> Opinion Pool questions #2 and #3 .....	<b>156</b>
<b>Graphic 7.</b> Opinion Pool questions #4 and #5 .....	<b>156</b>
<b>Graphic 8.</b> Opinion Pool questions #6 and #7 .....	<b>157</b>
<b>Graphic 9.</b> Opinion Pool questions #8 and #9 .....	<b>157</b>
<b>Graphic 10.</b> Overall mean results from the opinion pool.....	<b>158</b>

# Chapter 1

## Introduction

Cooperation and coordination are intrinsic characteristics of musical practice. Independently of the task performed in the music creation process (composition, performance, improvisation teaching or learning) individuals are inevitably confronted with bilateral or collective collaborative scenarios.

Likewise, computer networks are based on collaboration. Since the Early 70's there is an underlying awareness in Western culture about the advances brought by Computer Science and Digital Communication to collaborative processes.

Up until the early 90's systems that approached collaboration mediated by digital technology were mostly based on local computer networks due to technical constrains in Electronics and Telecommunications. However, recent technological advances, particularly in Internet computing, made available to the common computer user different types of collaborative tools, such as simple e-mail systems, textual chats, shared editors, video conference systems or shared spaces for the exchange of multimedia documents.

Collaborative Work performed by geographically dispersed individuals became a research field of the most importance in modern information society. And even though there is still some imprecision about the exact focus of this field<sup>1</sup>, the advent of group synergy maximization, in terms of time and space, and its impact on teams and organizations' workplace settings, can be regarded as a primary motivation behind the sudden growth of this area.

---

<sup>1</sup> Similar sounding terms, as workgroup computing, collaborative computing, groupware, cooperative work support, are constantly coming up to characterize this field of research and development.

In the same way, it is also a common practice for artists to use cutting edge technology in order to maximize the aesthetics and conceptual value of their work. This is generally achieved by enhancing the efficiency of creative processes and by using technology as a media itself to express meaningful artistic work, attempting to achieve stylistic and conceptual originality.

Inevitably, the development of systems based on computer networks for musical practices, emerged as a natural development that dates back to the late 1970's with experimental musical performances by the League of Automatic Music Composers<sup>2</sup>.

In addition, the massive world wide growth of the internet network is characterized by a community of users strongly moved by music in many different ways. Today we face a new medium of acoustic collaboration with a shared nature, which offers new prospects for music creation.

Music Technology is a Research field inherently appropriate to provide a context of study in this area. The Music Technology Group at the Pompeu Fabra University in Barcelona, founded by Xavier Serra in 1994, as well as the hospitable and open atmosphere of the Interactive Systems Group leaded by Sergi Jordà, was ideal for the gestation of the research work accomplished in this doctorate thesis.

## 1.1 Motivation

Computer-Supported Cooperative Work for Music Applications is an open area of research with many questions to be addressed, such as:

- What is the role of these systems' creators?
- Is community music an aesthetical meaningful Sonic Art form?
- How do the sonic results of these systems fit into time and space as we know it in the musical context?

---

<sup>2</sup> Early Experiments with musical networks are discussed in Chapter 2, Section 2.3.1, of this dissertation.

- Is the regular internet user ready to go beyond the role of a spectator and become a creator, composing, performing or improvising in a music piece?
- What kind of constraints can, or should, be considered in the user interaction layer?
- Is Multimodal interface design a possible approach for community music?
- Should one aim for a visual environment interaction models that are driven by the Soundscape?
- What should be handled by the user and what should be handled by the system?
- Could there be defined general-purpose models at the architectural and acoustic communication level that address the specificities of this paradigm?

## 1.2 Objective of this Dissertation

The objective of this thesis is to contribute to the field of Computer-Supported Cooperative Work for Music Application, building upon the hypothesis that new collaborative paradigms are bounded to the unique characteristics introduced by the use of computer networks in the mediation process of musical applications.

More specifically, it is focused on:

- Study and research of music interaction models and how they can be adapted to the unique facets of a global computer network;
- Methods to overcome, or diminish, the disrupting effects of network features in acoustic communication for musical practice.

## 1.3 Structure of this Dissertation

This doctorate work followed a methodology that departed from a contextualization and survey of the field, followed by a definition of concepts and ideas which represent advances in the field and concluded by a test-proof implementation of a software prototype and respective evaluation.



Three distinct parts can be defined within the main contents of this dissertation:

**Part I (Chapter 2)**

Systematic study and classification of state of the art systems for computer-supported cooperative work, with particular emphasis on geographically displaced musical practices.

**Part II (Chapters 3 and 4)**

Detailed analysis and discussion of experimental proposals, concepts and methods, based upon the contextualization studies.

**Part III (Chapter 5)**

Discussion about the implementation and evaluation of a proof-of-concept software prototype entitled Public Sound Objects.

# Chapter 2

## Survey of Computer-Supported Cooperative Work for Music Applications

Technological innovation has always contributed to the evolution of musical creation, leading to the construction of new instruments and thus offering greater possibilities of composition, interpretation and performance practice as well as facilitating the emergence of new sounds and stylistic innovations.

Likewise, computer networks empower cooperation between musicians, through the interconnection of electronic devices and the possibilities offered by synchronous and asynchronous exchange of information. We are then facing a new medium of musical communication with its own specific characteristics and new prospects of creation.

This topic has been addressed by the Computer Science Community, outside the scope of music creation, through an area of study entitled Computer-Supported Cooperative Work (CSCW)

### 2.1 Computer-Supported Cooperative Work

CSCW is acknowledged as a very significant research field representing one of the major focus areas of the former Special Interest Group on Groupware (SIGGROUP), from the world's first educational and scientific computing society - Association for Computing Machinery Organization (ACM). SIGGROUP was formally dissolved in January 30, 2005 by ACM, since it was considered that the CSCW and GROUP conferences were both technically and financially healthy, and they would only require the oversight of ACM and other SIGs.

The term CSCW was introduced by the computer scientists Irene Greif of Massachusetts Institute of Technology (MIT) and Paul Cashman of Digital in the early eighties.

According to the article written in 1992 by Liam J. Bannon's about CSCW (Bannon, J., 1992), Cashman and Greif came up with the term Computer-Supported Cooperative Work to describe the object of interest for a small workshop they organized in Massachusetts (August 1984), concerning the development of computer systems that would support people in their work activities. The workshop brought together people from different areas, such as office information systems, hypertext and computer-mediated communication.

Since the first Computer-Supported Collaborative Work conference organized in December 1986 in Austin, Texas, the enthusiasm for the topic continued to grow until today, with an increasing activity in the research and development of systems and applications as well as the publication of research work.

Besides the major CSCW conferences, in recent years there have been a number of CSCW-related conferences and workshops on collaboration technology, group decision support systems and multi-user systems both in Europe and North America, and in addition, several journals include CSCW in their list of topics.

Finding a commonly accepted definition of CSCW and its scope has been a difficult task mostly due to the multidisciplinary nature of the field which brings together people across a range of different backgrounds like computer science, psychology, sociology, organizational theory, and anthropology, just to mention a few.

A very general notion can be defined in the sense that "it is focused on the design of computer-based technologies with explicit concern for the socially organized practices of their intended users" (Suchman, L., 1989) (Bannon, J., 1994), but when designing specific applications to suit the specificities of certain work contexts, major differences between the resulting solutions might come along.

It is therefore not surprising that different tendencies emerged in the field. On one hand, there are groups focused on modeling and designing office communication systems. On the other hand there are those interested in developing a richer understanding of cooperative work practices.

This last approach is also followed in even more specific areas of interest, oriented towards general artistic creation and in particular to musical and sonic expression.

### 2.1.1 Operation Modes in CSCW

In general a group of users operating in joint projects will follow a specific approach to their contribution according to the project requirements or technological constraints of the system being used.

In a 1991 survey of CSCW systems (Rodden, T., 1991), Tom Rodden presents a systematic approach defining different operation modes in CSCW starting by classifying existing systems/applications and identifying their functional roles:

**Message systems** – In these systems the users operate in Transference mode interchanging information documents, but the work development is done individually, without a common sense of the global information structure (e.g. e-mail based systems).

**Computer Conferencing** – Information regarding a certain topic is broadcast towards an interested community. All the users participate and cooperate at the same level in the joint event and information is normally held with conference messages within one central database rather than the individual mailbox approach used in messaging systems. The development of reliable high speed communications has led to the emergence of new real-time conferencing systems, allowing conference members to communicate in real-time, and enhanced the scope and power of this class of applications (e.g. video/audio conference; news groups).

**Coordination Systems** – Addresses the problem of integrating and adjusting in a harmonious fashion the synergies of a group of people working together in the same physical space, by introducing the support of computer systems (e.g. white boards; automated meeting rooms with a network structure to support voting systems; multi-user software based on analytical decision techniques).

**Co-Authoring Systems** – General class of systems that supports the co-authoring of a product, designed to address the specificities and requirements of the product following a structured development of the content (e.g. Systems for joint development of software, with characteristics like maintaining up-to-date versions of the code produced by each project member and integration mechanisms for their partial contributions).

Even though this classification is quite accurate in the context of office communication systems, when considering an application out of this scope, like a scenario where a computer network is used for a collective artistic creation, it is not totally applicable<sup>3</sup>.

On the other hand, Rodden also presents in his survey more general characteristics common to CSCW systems, which could be considered as the environmental facets of cooperative work. The geographical nature of the user is considered as the Space Dimension environmental facets and can be “Remote” or “Co-located”.

The form of interaction provides the Time Dimension environmental facets, and it can be “Synchronous” or “Asynchronous”.

The aspects of synchronicity are extremely relevant when characterizing the operation mode of a joint system and it is a topic that requires special attention.

## 2.1.2 Synchronous and Asynchronous Modes in CSCW

The users operation in a joint project can be performed in a synchronous or asynchronous mode.

In the synchronous mode, all the participants are active simultaneously on the common document.

In the asynchronous mode, the participants do not need be active simultaneously, although the system must support the situation in which several participants happen to be active at the same time.

Practical experience with computer conferencing systems, which were initially designed to operate in both a synchronous and an asynchronous mode, tells us that almost all usage is in the asynchronous mode, showing that when both modes are available, users nearly always choose the asynchronous mode for serious interchanges.

In fact, the advantages of the asynchronous mode are often quoted as important reasons for using computer conferencing systems:

---

<sup>3</sup> The classification of different systems for collective artistic creation in the context of music and sonic arts is discussed in this chapter in section 2.4.

- All participants need not find a time in which they can all be active simultaneously. If one participant cannot participate at a certain time, that participant can enter the discussion a few days later and will still be able to contribute to the outcome of the work.
- You are not forced into rush decisions because of the time limits of a simultaneous meeting. If, for example, there is a need to think more about a problem, collect facts or make tests, this can be done and new input to the decision can be given the next day, without delaying the whole decision until the next time when all are together simultaneously.
- Some people need more time than others to read and reflect on a problem, because of the asynchronous nature of the system each participant can choose to spend more or less time on the topics of discussion as needed.

It is legitimate to assume that the same reasons that lead people to prefer an asynchronous mode of operation in computer conferencing are equally valid for joint editing systems, and therefore, these should be designed to work well in asynchronous circumstances.

Another point that is necessary to clarify is that although many times the concept of synchronous collaboration is referred to as being bound to real-time collaboration, in many situations, it is not necessarily so.

One possible example is a scenario where different artists are collaborating in a piece during an event with a limited duration in time. Even if the awareness of each artist regarding the contributions from the other is affected by latency impeding a reaction in real-time, the fact is that if this latency is much smaller than the piece duration, artists will be able to react to each other several times in a synchronous way during the course of the piece.

Usually the synchronization of the participants' contribution in joint remote performances for events is required, even if real-time communication is not supported as it would if the performance was based on physical presence.

### 2.1.3 The CSCW Classification Space

Based on the definition of CSCW environmental facets and the classification of existing systems, Rodden presents in his survey a classification space<sup>4</sup> graphically represented in the following figure.

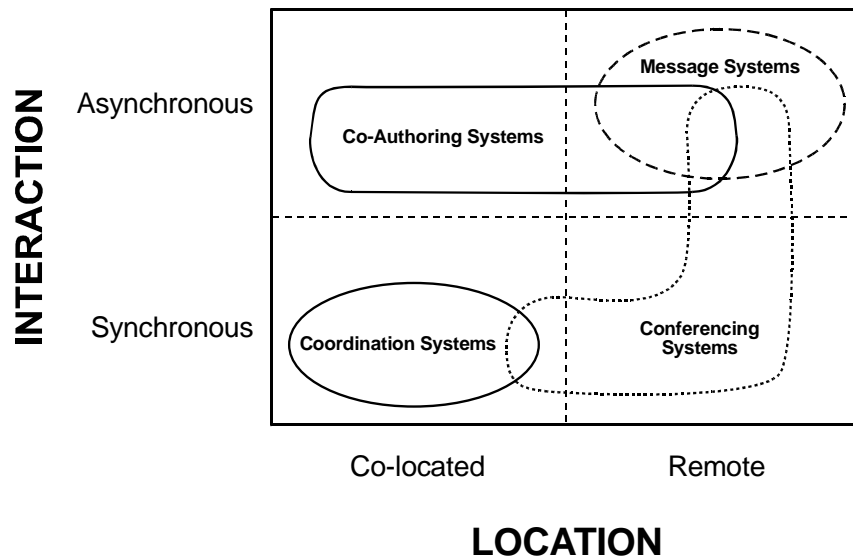


Figure 1. Rodden's Classification Space for CSCW Applications

A more extensive and detailed classification of web based collaborative systems was published in 2001 by Georgia Bafoutsou and Gregory Mentzas from the National Technical University of Athens, Greece at the 12th International Conference on Database and Expert Systems Applications (Bafoutsou, G. and Mentzas, G., 2001).

### 2.1.4 Shared Virtual Environments

Shared Virtual Environments (SVEs) go beyond the typical CSCW applications mentioned so far. These systems are not conceived with the main purpose of maximizing the synergies of a group in order to achieve better and faster results in the common tasks performed by the users.

---

<sup>4</sup> Rodden's original classification space, refers class of applications defined as Meeting Rooms. Meeting Rooms are a subset of the general class of applications defined here as coordination systems.

In a SVE it is intended to create a shared space in a computer network, inside which users can achieve a certain degree of immersion and flexibility in their behavior. This virtual space provides a context in which the interaction outcome is somehow unique and highly influenced by the characteristics of this media. Even though SVEs are more about simulations and experimentation than improving the efficiency of specific tasks, the flexibility and scalability of these systems often have proven to be effective when applied to contexts like education support or conferencing.

The pioneer internet systems that convey the essence of a virtual community space started with the original Multiple-User Domain/Dungeon (MUD) software developed in the early 80's by Richard Bartle and Roy Trubshaw at the University of Essex in England. A MUD is a real-time structured textual chat forum. It has multiple "locations" like an adventure game, and may include combat, traps, puzzles, magic, a simple economic system, and the capability for characters to build more structure onto the database that represents the existing world.

In the late 80's students on the European academic networks quickly improved the MUD concept, creating several new MUDs (VAXMUD, AberMUD, LPMUD). In many of these systems research was done to include bulletin-boards and social interaction mechanisms which added academic value to the projects. This, along with the fact that Usenet feeds were often unreliable and difficult to get in the U.K., made the MUDs a major form of social interaction in the online community at this time.

By 1996 Pavel Curtis from Xerox Corporation, introduced the cutting edge of this technology with the Object-Oriented MUD (MOO), an even more extensible system, using a built-in object-oriented language, allowing greater programmability and flexibility.

In a MUD or in its successors like the MOO or the Internet Relay Chat (IRC), the participants, usually called players, tend to overcome the constraints imposed by a text-based form of communication developing a specific language for communication and collaboration amongst themselves, that evolves to some sort of virtual social behavior, that only makes sense in these environments (Marvin, L. E., 1995).

In his article from 1992 "Mudding: Social Phenomena in Text-Based Virtual Realities", Pavel Curtis discusses how the emergence of MUDs created a new kind of social sphere and clarifies critical notions that tend to lead to misunderstandings about this paradigm, like the idea that since a MUD is somehow linked to entertainment it is also a computer game



“...Three major factors distinguish a MUD from an Adventure-style computer game, though:

(1) A MUD is not goal-oriented; it has no beginning or end, no ‘score’, and no notion of ‘winning’ or ‘success’. In short, even though users of MUDs are commonly called players, a MUD isn’t really a game at all;

(2) A MUD is extensible from within; a user can add new objects to the database such as rooms, exits, ‘things’, and notes. Certain MUDs, including the one I run, even support an embedded programming language in which a user can describe whole new kinds of behavior for the objects they create;

(3) A MUD generally has more than one user connected at a time. All of the connected users are browsing and manipulating the same database and can encounter the new objects created by others. The multiple users on a MUD can communicate with each other in real time. This last factor has a profound effect on the ways in which users interact with the system; it transforms the activity from a solitary one into a social one.” (Curtis, P., 1992)

Parallel to this text based systems, which are highly accepted by the on-line community up until today, other approaches emerged closer to the fields of Virtual Reality. The first large scale commercial networked multi-user virtual world was called **Habitat**<sup>5</sup> and it was developed by Lucas-Films Games in association with Quantum Computer Services, Inc in 1985. It worked on Commodore 64 computers, and ran for 6 years in Japan and the US.

Systems like **Active Worlds**<sup>6</sup> or **DIVE**<sup>7</sup>, which are highly sophisticated distributed interactive 3D virtual environments, came up in the early 90’s. These systems took advantage of the

---

<sup>5</sup> Habitat was conceptually a game. Yet, it introduced the concept of **Avatars** (the incarnation of a user within the system) and collaborative interaction within the virtual environment, a concept which is the basis of today’s Virtual Worlds.

<sup>6</sup> Active Worlds is a comprehensive platform for delivering real-time dynamic and visually compelling interactive 3D content over the web. It was a pioneer system of its kind since 1995, and in one of its instances “The AlphaWorld” there are 200,000 users registered.

combination between internet computing and the increasing computing power in graphic processing.

The idea of impersonating real life in a virtual 3D world, in which users can participate representing themselves as a 3D Avatar and achieving a high degree of immersion, seems to be the basis of these paradigms. Nevertheless it is surprising how projects that have a much simpler approach in terms of the realistic reproduction of real life with computer graphics, like the **Habbo Hotel**<sup>8</sup> launched in January 2001, tend to have excellent acceptance amongst the online-community.

## 2.2 Computer Mediated Communication and Networked Music

“What I want to say about Networked Music in general is that All Music Is Networked. You can think about an Orchestra as client-server network, where a conductor is “serving” visual information to the “client” musicians, or a peer-to-peer networking model in an improvising Jazz Combo, where there is no one directing, and the musicians are all interacting, so, any performance context we can think of in some way there is a network connecting the performers (...). Networked Music with capital N and capital M (the kind we are talking about) is about performance situations where traditional aural and visual connections between participants are augmented, mediated or replaced by electronically-controlled connections.” (From Jason Freeman’s lecture opening at the 1st Networked

---

<sup>7</sup> DIVE (<http://www.sics.se/dce/dive/>) stands for Distributed Interactive Virtual Environment and it is a non-commercial experimental system developed by the Distributed Collaborative Environments Group from the Swedish Institute of Computer Science.

<sup>8</sup> The Habbo Hotel (<http://www.habbohotel.com/>) Habbo Hotel is a graphical chat environment for the Internet, holding a community of nearly two million members. Built in Macromedia Shockwave, the website takes the form of a virtual 3D hotel displayed in 2D Graphics.

Music Workshop during the International Computer Music Conference  
2005)

From his definition of Networked Music Jason Freeman clearly illustrates why collaborative music paradigms can easily be approached in the context of computer mediated communication.

### 2.2.1 Collaboration in Music from a psycho-social perspective

Music has always been part of social life, both as a mean of cultural expression and as a mechanism to enhance social cohesion. Present in many of the collective and ceremonial rituals since early times and across all cultures, collaboration and coordination among performers are intrinsic characteristics of collective musical creation.

On the other hand computer mediated communication and the ubiquitous nature of the Internet brought out a new facet to the community, which also applies to the context of sonic and musical arts. This new facet incorporates both the emergence of new types of musical instruments and the consequent re-definition of the individuals' functional roles in music creation (sharing the creator, performer and listener' roles). Furthermore it presents the possibility for both an individual and a group to experiment music creation and collaboration laid on a very abstract language that is, nonetheless, accessible to all.

From a perspective of social psychology of music, Hargreaves suggests that the music psychologist's job is to investigate the multiple ways in which we engage in music and try to explain the mechanisms that influence our behavior (Hargreaves, D., Miell, D. and MacDonald, R., 2005). According to these authors, this musical behavior must be investigated in all of the social and cultural contexts in which music takes place, which has been largely widened by the widespread availability of the Internet. Therefore this field of research is taken beyond specific, formal 'musical' research scenarios.

This is the reason why the social psychological approach calls for a greater emphasis on the study of musical behavior in everyday life situations. In fact, music is essentially a social activity, and it is commonly accepted that the social functions of music will have an impact on the individual's social activity, cognition and emotion.

In this context of both geographical and social dispersion, individuals collaborating amongst themselves become **virtual communities**. According to Bogazzi & Dholakia (Bogazzi, R. P. and Dholakia, U. M., 2002), a virtual community is a mediated social space in a computer

mediated environment, which allows the establishment of groups. It is basically supported by communication processes of continuity. These spaces are currently created by means of a diverse set of technological possibilities previously discussed in section 2.1 of this chapter.

Virtual Communities are normally formed on the basis of a specific interest shared by its members, which is the reason for their affiliation. Each person thus gains a sense of affinity with the rest of the community members as well as a sense of collective identity that distinguishes him from non-members. So, more than the majority of communities established through face-to-face interaction, virtual communities result from a conscious choice of each member to participate. A final inherent characteristic of virtual communities is the fact that individuals are creators and not consumers. Virtual Communities' members often are highly specialized in the specific subject conveyed even in the most common community activities (Bogazzi, R. P. and Dholakia, U. M., 2002).

Nevertheless, Virtual Communities also present differentiated characteristics amongst them:

- Asynchronous communication (as in mailing lists), synchronous (as in chat-rooms) or both;
- Verbal language (still the majority of the cases), visual language (static or moving) or acoustic language;
- Open participation or participation subject to pre-defined conditions;
- Functional (useful to the lives of its participants) or hedonistic goals (based on the pleasure derived from communication and content creation).

One can also consider the level of the community's internal structure.

- A high structural level creates in the communities a strong interdependence among its members, originating phenomena typical of the processes of group formation: the development of sanction norms and its mechanisms, affective bonds or group identity.
- A low structural level is associated to anonymity conditions that may lead to the reduction of the sense of individual responsibility concerning the final results. This way, there would be less perceived social pressure, which may well contribute to a greater experimental and creative freedom.

All the characteristics and constraints discussed in the context of Virtual Communities also apply to a scenario of acoustic communication over the internet, or through the exchange of control data between virtual Music Instruments<sup>9</sup>.

### 2.2.1.1 Towards a New Social Space of Music creation

In 1984 Barry Truax introduces the concept of Acoustic Community at the book “Acoustic Communication” (pages 57 and 58):

“Thus far we have concentrated on a model of acoustic communication from the perspective of the listener in which listening is understood as the primary acoustic interface between the individual and the environment. However, the flow of communication goes both ways since the listener is also a sound maker, and therefore it is the entire system of the listener plus the environment which constitutes the ‘Soundscape’. (...) The ‘Acoustic Community’ may be defined as any Soundscape in which acoustic information plays a pervasive role in the lives of the inhabitants (...). Therefore the boundary of the community is arbitrary and may be as small as a room of people, a home or a building, or as large as an urban community, a broadcast area or any other system of electroacoustic communication.” (Truax, B., 1984)

The notion of Soundscape addressed by Truax, was extensively discussed in 1977 by Murray Schafer in the book “The Soundscape – Our Sonic environment and the turning of the world”. Schafer defines Soundscape as:

“The Sonic Environment. Technically, any portion of the sonic environment regarded as a field for study. The term may refer to actual environments, or abstract constructions such as musical compositions or tape montages, particularly when considered as an environment.” (Schafer, M., 1977)

---

<sup>9</sup> The concept of Virtual Music Instruments and its implications when applied over a network of computers is discussed in Chapter 3, Section 3.2.1.

Truax's concept of acoustic community, regarded in the context of a virtual community, conveys the notion of a Soundscape common to users displaced in time and space, and is inevitably tied to sonic events perceived and produced through computers. It is the paradigm of a **Shared Soundscape**.

The influence of the Shared Soundscape paradigm, as any new paradigm of music creation, may well have an effect upon the individual's creation abilities. In fact, according to Csikszentmihalyi's view of the phenomenon of creativity (Csikszentmihalyi, M., 1999), creativity is much a psychological as it is a cultural and social event and what we call creativity is not the product of single individuals, but of social systems making judgements about individuals products. To develop this perspective, he used a 'systems' model of the creative process that takes into account these three essential features. In sum, it is a theoretical framework that aims to explain the relationship established between the creative individuals and their socio-cultural contexts, organized around the following analytical three axes: **Domain** (a cultural or symbolic aspect), **Field** (a social aspects) and the **Creative Person**.

- The Domain is constituted by the accumulated knowledge in the area and is operated by means of a set of objects and tools, representation rules and notations.
- The Field comprises specialists, professionals or those who judge new Works and influence the way in which the works are socially accepted or rejected. Therefore, their actions build the consensus inter-subjectively at a given moment.
- The creative individuals are those who transform the fields in which they act. There are several conditions that favor innovative action, such as personal characteristics, the dedication to experimentation or a privileged position in the domain.

Csikszentmihalyi then defines creativity as a process that can be observed only at the intersection where individuals, domains and fields interact. In the following section I will describe this system, by using the example of the Shared Soundscapes paradigm so as to show that both this new domain and the individuals' who act upon it, introduce a redefinition of several aspects of the music making activity that also allow the emergence of creativity.

### 2.2.1.2 Csikszentmihalyi's Creative Person in Shared Soundscapes

In Shared Soundscapes there are two categories of creative people: experts and non-experts. On one hand, there are the experts in computer technology, who lead innovative creation. In order to play this role they are required to master musical and technical skills.

This qualification can be acquired either through the formal educational process or through a self-learning process. Those who gained their expertise through formal education are generally Computer Science experts with musical knowledge or the reverse.

On the other hand, non-expert users, with no formal education, also participate actively in virtual communities, following a profile of familiarity with Computer Technology and urge to explore the possibilities provided by global connectivity through the Internet.

Naturally, as in any aesthetic language, in order for the works to be meaningful, it is necessary to have a certain degree of acquired knowledge and familiarity. However, interfaces and interaction paradigms are increasingly tailored to suit an average Internet and Computer System user. Since these musical languages are not necessarily based on a traditional musical grammar, its control is acquired by a trial and error approach and consequent adjustments made by the different participants.

These creative communities, by their nature, seem to assume hedonistic and not functional objectives. As no material achievement or utilitarian aims are fulfilled in such joint creations, it is the search of aesthetic and creative pleasure per se, which results from the personal identification with this type of activity that forms the basis of the motivational process. The challenges taken are the exploration of the language itself and the coordinated interaction with other participants.

### 2.2.1.3 Csikszentmihalyi's Domain of Shared Soundscapes

Taking into account the theoretical construct of Csikszentmihalyi (Csikszentmihalyi, M., 1998), In Shared Soundscapes there are two categories of creative people: experts and non-experts. On one hand, there are the experts in computer technology, who lead innovative creation. In order to play this role they are required to master musical and technical skills.

This qualification can be acquired either through the formal educational process or through a self-learning process. Those who gained their expertise through formal education are generally Computer Science experts with musical knowledge or the reverse.

On the other hand, non-expert users, with no formal education, also participate actively in virtual communities, following a profile of familiarity with Computer Technology and urge to explore the possibilities provided by global connectivity through the Internet.

Naturally, as in any aesthetic language, in order for the works to be meaningful, it is necessary to have a certain degree of acquired knowledge and familiarity. However, interfaces and interaction paradigms are increasingly tailored to suit an average Internet and Computer System user. Since these musical languages are not necessarily based on a traditional musical grammar, its control is acquired by a trial and error approach and consequent adjustments made by the different participants.

These creative communities, by their nature, seem to assume hedonistic and not functional objectives. As no material achievement or utilitarian aims are fulfilled in such joint creations, it is the search of aesthetic and creative pleasure per se, which results from the personal identification with this type of activity that forms the basis of the motivational process. The challenges taken are the exploration of the language itself and the coordinated interaction with other participants.

#### 2.2.1.4 Csikszentmihalyi's Field in Shared Soundscapes

Despite its evident connection with musical expression, the language and the instrument used in Shared Soundscapes are enclosed into the field of Sonic Arts discussed in the following section of this chapter,.

The Aesthetic Sonic Language is the outcome of two input streams. The first is the high level conceptualization and design of the system. The second is the low-level programming of the system, which is somehow invisible to the end user, but many times determines the interaction process and the system efficiency. Its development is highly dependent on the computational platform and programming language, in which the system is to be implemented.

These two streams in the implementation process often result in a collaborative process between creators from different Domains of expertise, even if their roles are clearly differentiated. The system creators generate a grammar, which the user will operate through an interface in order to achieve the final outcome of a Sonic Art Piece.

In terms of aesthetic languages these computational systems, when flexibly conceived, allow the users to choose, both an approach tied to musical tradition, or a more experimental and free



form of acoustic expression. Shared Soundscapes is a highly experimental field in which each new system is likely to represent an innovative proposal to a language that is not fully established yet.

## 2.2.2 Redefining the Acoustic Community for Music and Sonic Arts

The notion that community space needs to be redefined for the internet community has also been stated in 2001 in the article “Observations about Music and Decentralized Environments”, in which Dante Tanzi points out:

“The placing of musical phenomena in the community space of the Net produces a change in the condition of appearance of musical events and their objectification. Therefore when faced with the plurality of the musical processes each of which has opened numerous references, it is advisable to agree on the criteria of recognizability of musical events within a different communication space”. (Tanzi, D., 2001)

What Tanzi suggests is that from the proliferation of systems which produce some sort of sonic outcome from the internet’s acoustic community, will result the possibility to acknowledge these results as musical events given that eventually they will be recognized as so. However, past traditional musical culture is somehow strict as to what is recognizable as a music event (even though this is a result of cultural tradition that can change in the future) and one of the major questions regarding collective music creation by indiscriminate Internet users is if this community is prepared to express meaningful expressive musical performances.

Given that music universe’s subset considered in this context is constrained to music produced on a computer network based interactive paradigm, it is useful to consider the definition of music by Guy E. Garnett in the Article “The Aesthetics of Interactive Computer Music” published in the spring of 2001 by MIT Press in the *Computer Music Journal*:

“The nature of music, particularly in the century of John Cage, multiculturalism, and other varieties of aesthetic choice, become more problematic. Nonetheless, I think it is possible to reduce the problem somewhat. Just as I have considered aesthetics in only its broadest manifestations, similarly, music can be roughly considered to be sounds

made with aesthetical intent or even sound listened to with aesthetic interest. The former gives more weight to the role of the creator, while the latter formulation tends to privilege the listener.” (Garnett, G., 2001)

In the broadest sense we could certainly consider that the sonic outcome of a collective internet creation is a musical event, since users commit themselves to collaboratively create a sound piece with an aesthetic intent and simultaneously are an audience (even though the audience could transcend the creators) interpreting the results from each other with an aesthetical interest.

In order to better define the subset of the music universe in which this approach for sonic creation is situated, the term Sonic Art emerged in Artistic communities since the 60’s. Historically, Sonic Art derives from the academic tradition of electroacoustic (electronic) music, since until quite recently, advanced electronic and computer technology for audio work has only been available to members of institutions such as universities and radio stations. This tradition dates back to the 1950’s and 60’s, when electroacoustic music discipline emerged in colleges and university music departments, based on the work of composers like Pierre Schaeffer and Karlheinz Stockhausen.

Even though there is no comprehensive definition of Sonic Art, with the advent of computer technologies to the common music creator in the 80’s, and with computer communication over the internet in the 90’s, this field became the playground for diversified artistic proposals and experiments for music creation with electronic and digital technology.

### 2.2.3 Networked Music as a Research Topic

The issue of whether Music Technology itself can be considered as an established research field has been raised by Xavier Serra in different occasions. From his experience as a researcher and director of the Music Technology Group in Barcelona, Serra analyzed the present situation in this area of work and proposed some preliminary ideas to design a Roadmap for research in Music Technology (Serra, X., 2005). Music Technology is a Multidisciplinary field involving many disciplines such as Music, Computer Science, Psychology, Engineering and Physics

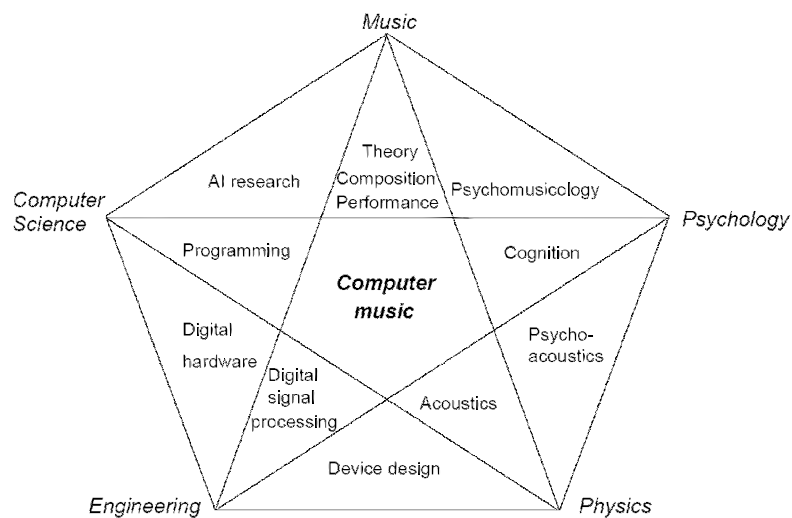


Figure 2. Diagram of the Computer Music interdisciplinarity field proposed by (Moore, F. R., 1990).

According to Serra, the fact that there are several conferences focused on the area of Music Technology, as well as a number of peer reviewed Journals from reference publishers in the Scientific Community, are strong indicators of good scientific dissemination.

In fact, most of the work related to music creation mediated by computer networks has been published in these conferences or Journals during the last decades<sup>10</sup>.

Main Examples of Music Technology conferences are:

International Computer Music Conference (ICMC); Digital Audio Effects Conference (DAFX); International Conference on Music Information Retrieval (ISMIR); New Interfaces for Musical Expression Conference (NIME).

Main Examples of Music Technology scientific journals are:

Leonardo Music Journal and Computer Music Journal from the MIT Press; Organised Sound Journal from Cambridge University Press; Journal of New Music Research from Routledge.

In 2003 during the International Computer Music Conference, held in the National University of Singapore, some awareness was raised about this type of research burgeoning to become one of the acknowledged topics of Music Technology.

---

<sup>10</sup> Further reference to historical work in this field will be discussed in section 2.3 of this chapter

In his Keynote Speech from ICMC 2003 Roger Dannenberg mentioned “Networked Music” as one of the promising research topics and at least four papers (Barbosa, A., Kaltenbrunner, M. and Geiger, G., 2003), (Stelkens, J., 2003), (Hajdu, G., 2003) and (Obu, Y., Kato, T. and Yonekura, T., 2003) were centered on this topic, even though they were scattered over different panels, instead of one distinct session.

Since then the term Networked Music has become increasingly consensual in defining the area, and according to Jason Freeman’s definition: *it is about music practice situations where traditional aural and visual connections between participants are augmented, mediated or replaced by electronically-controlled connections.*

### 2.2.3.1 Landmarks in Networked Music Research

In order to have a broad view over the scientific dissemination of Networked Music research I present the following Landmarks in the field over the last six years (1999-2005):

(1) The ANET Summit; (2) The Networked Music Workshop at ICMC; (3) Four published Doctorate Dissertations; (4) Six surveys and partial overviews published in journals about Networked Music; (5) A dedicated issue to Networked Music from Cambridge Press’ Organised Sound Journal.

#### (1) The ANET Summit (August 20-24, 2004)

The summit was organized by Stanford University’s Center for Computer Research in Music and Acoustics (CCRMA) and held at the Banff Center in Canada, was the first Workshop event addressing the topic of High quality Audio over Computer Networks. The guest lecturers were **Chris Chafe, Jeremy Cooperstock, Theresa Leonard, Bob Moses and Wieslaw Woszczyk.**

The Workshop Syllabus stated:

“This three-day summit is an exploration of the state-of-the-art in ethernet-based professional audio networks. Developers, engineers, musicians and others interested in the growing practice of high-resolution audio over ethernets will gather to focus on the new technology. The scope includes IP-based systems and systems with dedicated protocols.

A 1998 AES whitepaper on "Networking Audio and Music Using Internet2 and Next-Generation Internet Capabilities" expressed a vision of the future

and challenges that lay ahead. Six years later, with technical developments continuing, musical collaborations of various kinds have been tested and the Internet has evolved. Predicted application areas which are now taking off include audio production, music education, broadening musical participation, and scientific and engineering data representation (sonification). The summit offers an opportunity to compare today's reality with what was foreseen and to look ahead to what's next.

The summit is a "neck-ties removed" working group that brings together academic and commercial interests, developers and users, audio specialists and network engineers. The program includes hands-on demonstrations in The Banff Centre's concert and recording facilities, a "how-to" covering representative open-source software-based systems, demos of products, presentations, a tutorial, and a panel discussion.

Continued topics from the 1998 vision of audio over next-generation networks include current and future quality of service, implications of end-to-end design, cost and complexity of bridge devices, formats and adherence to audio industry standards, and scalability requirements. New topics will include but are not limited to Internet signal processing, User studies, and new artistic forms." (Syllabus from the ANET Summit)



Figure 3. Participants of the ANET Summit; Organizers: Chris Chafe (1st left), Jeremy Cooperstock (4th Right), Theresa Leonard (6th Right), Bob Moses (3rd Right), Wieslaw Woszczyk (3rd left)

**(2) The Networked Music Workshop at ICMC (4th of September, 2005).**

This Workshop was held in Barcelona and resulted from experience in previous ICMCs, which called for the need to realize such an event. Guest Lecturers were: **Álvaro Barbosa** (Pompeu Fabra University, MTG), **Scot Gresham-Lancaster** (Cogswell College Sunnyvale, CA), **Jason Freeman** (Georgia Institute of Technology), **Ross Bencina** (Pompeu Fabra University, MTG).

The Workshop Syllabus stated:

“Participants in this workshop will learn about different types of networked music practice and about tools and techniques which are available to undertake these tasks. The focus will be on the technical, compositional, and aesthetic challenges involved in realizing networked music on local area networks and over the Internet, using both peer-to-peer and client-server networking models. The workshop will discuss systems intended to lead non-musicians towards creative expression, as well as systems for practicing musicians to extend the boundaries of performance.

After a broad overview of historical projects, like from the work of the League of Automatic Music Composer and the HUB from the 1970's and 1980's and other important approaches to networked music, the workshop will focus on case studies of particular projects and the tools they use. Ross Bencina's network transport infrastructure, OSCgroups, which uses a centralized name lookup server and peer-to-peer data interchange via Open Sound Control, will be discussed in connection with recent work by Scot Gresham-Lancaster and "The Hub." Phil Burk's TransJam, a Java-based server for real-time collaboration over the Internet, will be explored in connection with the Auracle networked sound instrument. Specific issues to be addressed include the logistics of event coordination, the mediation between transparency and complexity in the system, the handling of timing and latency issues, human interface design, and the maintenance and monitoring of client reliability.” (Syllabus from the Networked Music Workshop at ICMC 2005)

### (3) Published Doctorate Dissertations

- **2002 Golo Föllmer** “Musikmachen im Netz Elektronische, ästhetische und soziale Strukturen einer partizipativen Musik (Making Music on the Net, social and aesthetic structures in participative music)” – Martin Luther Universität Halle-Wittenberg – Germany.
- **2002 Nathan Schuett** “The Effects of Latency on Ensemble Performance” – Stanford University, California – USA.
- **2003 Jörg Stelkens** “Netzwerk-Synthesizer (Network Synthesizer)” – Ludwig Maximilians Universität, München – Germany.
- **2003 Gil Weinberg** “Interconnected Musical Networks – Bringing Expression and Thoughtfulness to Collaborative Music Making” - Massachusetts Institute of Technology, Massachusetts – USA.

### (4) Surveys and partial overviews published in journals about Networked Music

- **1999 Sergi Jordà**, “Faust Music On Line (FMOL): An approach to Real-time Collective Composition on the Internet”, Leonardo Music Journal, Volume 9, pp.5-12.
- **2001 Dante Tanzi**, “Observations about Music and Decentralized Environments”, Leonardo Music Journal, Volume 34, Issue 5, pp.431-436.
- **2002 Gil Weinberg**, “The Aesthetics, History, and Future Challenges of Interconnected Music Networks”, Proceedings of the International Computer Music Conference, pp.349-356.
- **2003, Álvaro Barbosa**, “Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation”, Leonardo Music Journal, Volume 13, Issue 1, pp.53-59.
- **2005, Gil Weinberg**, “Interconnected Musical Networks: Toward a Theoretical Framework”, Computer Music Journal, Vol. 29, Issue 2, pp.23-29.
- **2005, Peter Traub**, “*Sounding the Net: Recent Sonic Works for the Internet and Computer Networks*”, Contemporary Music Review, Vol. 24, No. 6, December 2005, pp. 459 – 481.

**(5) Organised Sound Journal Volume 10, Number 3 (December 2005) – Dedicated Issue to topic of Networked Music**

This Issue of Organised Sound, edited by Leigh Landy, is entirely dedicated to Networked Music.

The Call for contributions stated:

“Interconnection has always been a fundamental principle of music, prompting experimental artists to explore the implications of linking their computers together long before the Internet reached the public consciousness. As the Internet achieved critical mass over the past decade, networking technology took centre stage as the key to a vast new territory of possibility, facilitating remote participation, distributed processing, and redefinition of musical space and time. The Web emerged as a virtual venue for countless musical purposes, and as analog acoustics transformed to digital representations, packets of data carried by IP from one address to another became a modern metaphor for air molecules transmitting the tone of vibrating body to eardrum.

As with any new technology, applications of networking to music have evolved from naïve proofs-of-concept to more sophisticated projects, and we stand now at a point when 'internetworking' is taken for granted, novelty is expiring and artistic goals more often transcend technical considerations. From this vantage, the essential question is not how networking and music are combined, but why. What is the unique experience that can be created? Whose role can be empowered or transformed: composer, performer, audience? Where can sound come alive that it couldn't otherwise? Networked music can reinterpret traditional perspectives on stagecraft, ensemble, improvisation, instrumentation, and collaboration, or enable otherwise impractical relationships between controllers, sensors, processors, inputs, and outputs. The network can be an interface, a medium, an amplifier, a microphone, a mirror, a conduit, a cloud, or a heartbeat.

The network is all of us. Music is the sound we make. Listen...” (Call for Articles for Organised Sound 10.3)



This Issue featured articles from: Golo Föllmer, Matthew Wright; Mara Helmuth; Ajay Kapur, Ge Wang, Philip Davidson and Perry R. Cook; Jason Freeman, Kristjan Varnik, C. Ramakrishnan, Max Neuhaus, Phil Burk and David Birchfield; Álvaro Barbosa; Evandro Manara Miletto, Marcelo Soares Pimenta, Rosa Maria Vicari and Luciano Vargas Flores; Gil Weinberg; David Birchfield, David Lorig and Kelly Phillips

## 2.3 Systematic Study of Networked Music Systems

Over the last decades artists have used cutting edge computer technology to maximize the aesthetics and conceptual value of their work, not only by enhancing the way they traditionally create, but also by using technology as a media in itself to express meaningful artistic work.

The idea of using computer networks as an element in collective artistic creation and performance (or when both come together in improvisation) was no exception, since it provides a particularly engaging opportunity to achieve stylistic and conceptual originality.

On the other hand, identifying the influence of aesthetic and conceptual values in specific techniques related with the media or technology with which an art piece is created is often a hard and unclear task.

The Berlin based writer Florian Cramer, who has published in the area of code poetry, comparative studies in literature and art, refers to this precise task in the context of Combinatory Poetry and Literature on the Internet.

“(…)Although it is difficult to distinguish a combinatory literature from other forms of literature ever since linguistics defined language as a combinatory system itself, combinatory poetry nevertheless could be formally defined as a literature that openly exposes and addresses its combinatorics by changing and permuting its text according to fixed rules (…)” (Cramer, F., 2000)

One could equally argue that it is also hard to differentiate the artistic influence of a computer-supported collaborative system in the music creation process, but since collaboration in itself is part of the traditional music language, this new form of communication will influence the paradigm according to new rules.

These new rules will not only be based on the new possibilities of geographical displacement and asynchronous collaboration, but they will also be strongly dependent on technological constraints, thus defining a different stylistic and conceptual way to create music.

In this section we will discuss many of the systems and ideas that emerged in the context of Music and Sonic Arts collaboration over computer networks.

These systems somehow adapted to the rules imposed by the new media even if developers were not always totally aware of this fact, since most of the times they use methods of development based on a project oriented perspective, and not towards systematic research methods.

As we will see in practical terms these rules are dependent on the type of system and to what extent it uses the possibilities presented by large or small scale computer networks.

### 2.3.1 Early Experiments with Musical Networks

The idea of the communication media influencing musical practice is by no means new nor bounded to technology.

One of the most remarkable examples in western music of the media's influence in music performance leading to stylistic novelty is the Venetian polychoral music style. This is a type of music of the late Renaissance and early Baroque eras which involved spatially separate choirs singing in alternation. It represented a major stylistic shift from the prevailing polyphonic writing of the middle Renaissance, and was one of the major stylistic developments which led directly to the formation of what we now know as the Baroque style.

The style arose from the architectural peculiarities of the imposing Basilica San Marco di Venezia in Italy. Aware of the sound delay caused by the distance between opposing choir lofts, composers began to take advantage of that as a useful special effect. Since it was difficult to get widely separated choirs to sing the same music simultaneously (especially before modern techniques of conducting were developed), composers such as Adrian Willaert, the maestro *di cappella* of St. Mark's in the 1540s, solved the problem by writing antiphonal music where opposing choirs would sing successive, often contrasting phrases of music; the stereo effect proved to be popular, and soon other composers were imitating the idea, not only in St. Mark's but in other large cathedrals in Italy.

The idea of different groups singing in alternation gradually evolved into the *concertato* style, which in its different instrumental and vocal manifestations eventually led to such diverse musical ideas as the chorale *cantata*, the *concerto grosso*, and the *sonata* (Reese, G., 1954) (Bukofzer, M., 1947).

In an even more complex scenario one can consider the musical interaction between performers to the extent of interdependency, in which each musician can not only control his instrument, but somehow intentionally dictate or influence how the performance of other musicians evolves over time. This concept is known as Musical Interconnection and has been extensively studied by Gil Winberg in the context of his research work at MIT MediaLab (Weinberg, G., 2002).

Once again this idea is not bounded to technological mediation, and very early examples of Music performance as an Interdependent art form can be found in non-western music, namely in traditional Indonesian Gamelan music of Bali.

Gamelan is a tuned percussion ensemble/orchestra that follows a form of group interdependence based on the concept of Heterophony<sup>11</sup>. Through this music the Hindu-Balinese universe and its cyclic order is represented in a cyclic acoustic design as a multidimensional, idealized representation of cosmic balance.

Each Percussion instrument is specified to mark off established time intervals in a nuclear theme extended over a number of "bars" (almost invariably in 4/4 time), against which other instruments play a largely independent countermelody. Another group plays rhythmic paraphrases of this theme, and a fourth group fills out the texture with delicate rhythmic patterns, resulting in an evolving rhythmic experience dependent on the way each musician relates to the orchestra (Perlman, M., 2004).

Even though this notable examples show us that some of the present ideas in Music are inherited from a rich past culture, it was with the advent of electronics and computer technology that the concept of an Interconnected Music Network (IMNs) was taken further, allowing multiple ways of crossed control between performers and instruments.

---

<sup>11</sup> Heterophony is a kind of complex monophony (music with just one part, such as Gregorian chant), where there is only one melody, but multiple voices each of which play the melody differently.

An early example of group communication using electronic technology in the field of performing arts is John Cage's 1951 famous "Imaginary Landscapes No.4" for twelve radios played by 24 performers (Cage, J., 1961).

In this piece Cage unleashed the expressive potential of technology to enhance acoustic group interdependency by using the then recently invented commercial transistor radio as a musical instrument providing a sonic medium for collaboration, procedures and rules in his piece. The composition score indicated the exact tuning and volume settings for each performer but with no foreknowledge of what might be broadcast at any specific time, or whether a station even existed at any given dial setting.

The explorations of the transistor radio as an infrastructure for collaboration opened the door for other explorations with the electronic media, which were not necessarily based on external sound production.

In the late 1970's the commercialization of personal computers in the United States, allowing fine tune network topologies, enabled the first groups of experimental musicians to create musical computer networks at a local area scale.

In the mid-1970s, from the San Francisco Bay Area, emerged the first ensemble to investigate the unique potentials of computer networks as a medium for musical composition and performance entitled "**The League of Automatic Music Composers**" (Brown, C. and Bischoff, J., 2005) (Bischoff, J., Gold, R. and Horton, J., 1978) (Chadabe, J., 1997).

Originally the "League" came together through the mutual interest of Jim Horton, John Bischoff and Rich Gold, naming their new genre of musical performance "Network Computer Music".



Figure 4. The League of Automatic Music Composers (Perkis, Horton, and Bischoff, left to right) performing at Ft. Mason, San Francisco 1981. Photo: Peter Abramowitsch.

By networking their computers, each composition could send and receive data from the other compositions, and for the first time to create programmable and detailed musical interconnections.

Connections were made via the 8-bit parallel ports available on KIM's<sup>12</sup> edge connectors. In such a case, the program on the receiving end would either periodically check the port for new data or more casually retrieve whatever data was there when it looked. At other times the connection was made via the KIM's interrupt lines which enabled an instantaneous response as one player could "interrupt" another player and send a burst of musical data which could be implemented by the receiving program immediately.

---

<sup>12</sup> The KIM-1 is the first computer developed by Commodore in 1976. The KIM-1 has 1152 bytes of RAM, 2048 bytes of ROM and 30 I/O-lines. Some of these lines are used to drive six 7-segment LED-displays and others are used to read the little hexadecimal keyboard.

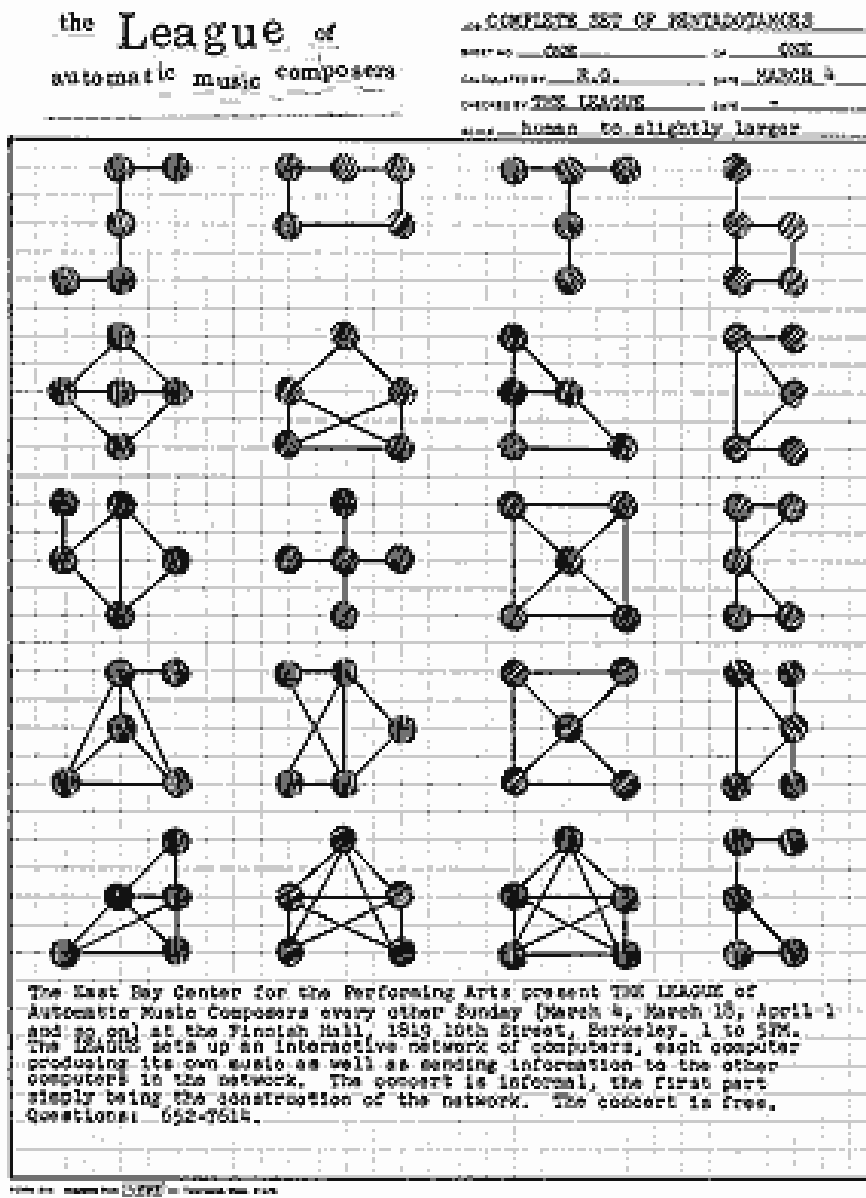


Figure 5. Flyer, designed by Rich Gold, from 1979 announcing a regular series of concerts, and showing different network connection topologies between the League computers.

"The League of Automatic Music Composers COMPLETE SET OF PENTADOTAMOEES The East Bay Center for the Performing Arts present THE LEAGUE of Automatic Music Composers every other Sunday (March 4, March 18, April 1 and so on) at the Finnish Hall, 1819-10th Street, Berkeley. 1 to 5 PM. The LEAGUE sets up an interactive network of computers, each computer producing its own music as well as sending information to the other computers in the network. The concert is informal, the first part simply being the construction of the network. The concert is free,

the first part simply being the construction of the network. The concert is free." (Flyer text)

In 1985 a Network music festival entitled "The Network Muse" was held in San Francisco, featuring a collective of electronic musicians including John Bischoff, Tim Perkis, Chris Brown, Mark Trayle, Scot Gresham-Lancaster, Larry Polanski, Phil Stone and Phil Burk.

From the context created around the activity of these composers interested in this new paradigm, "The League of Automatic Music Composers" evolved into a subsequent group, "**The Hub**" which employed more accurate communication schemes by using the Midi protocol.

In 1987 composers Nick Collins and Phill Niblock invited members of the Hub to create a performance that would link two performance spaces, Experimental Media and The Clocktower in New York City, to exemplify the potential of network music performance to link performances at a distance.

Two trios performed together in each space, each networked locally, and communicating with each other automatically via a modem over a phone line. This performance known as "**The Clocktower Concert**" was the first concert of the Hub, and a milestone in Networked Music by the incorporation of geographical displacement between performative Spaces.

### 2.3.2 Geographical Displacement in Music Communication

With the advent of global communication in the internet era, breakthrough possibilities to provide acoustic connections between worldwide displaced creators enhanced tremendously the traditional collaboration paradigm in Music and Sonic Arts.

The Internet brought an exponential increase of different possible scenarios in which collaborative music practice became possible, and a first systematic classification of different Networked Music Systems was published by Gil Weinberg in (Weinberg, G., 2002) and later in (Weinberg, G., 2005).

From his studies on Interconnected Music Systems Weinberg proposes four different approaches which characterize different branches of musical interaction, which differ in the level of interconnectivity among players and the role of the computer in enhancing interdependent social relations:

**The Server approach** - This simple approach uses the network merely as a means to send musical data to disconnected participants and does not take advantage of the opportunity to interconnect and communicate between players. Participants in such a server/client configuration cannot listen to, or interact with, their peers and the musical activities are limited to the communication between each player and the central system.

**The Bridge approach** - The motivation behind the Bridge approach is to connect distanced players so that they could play and improvise as if they were in the same space. Unlike the Server approach, musical collaboration can occur in such networks since participants can listen and respond to each other while playing. However, the role of the network in this approach is not to enhance and enrich collaboration, but to provide a technical solution for imitating traditional group collaboration. Aspects of bandwidth, simultaneity, synchronization, impact on host computer, and scalability are some of the challenges that are usually addressed in this approach.

**The Shaper approach** - In the Shaper approach the network's central system takes a more active musical role by algorithmically generating musical materials and allowing participants to collaboratively modify and shape these materials. Although players in Shaper networks can continuously listen and respond to the music that is modified by all participants, the approach does not support direct algorithmic interdependencies between players.

**The Construction Kit approach** - This approach offers higher levels of interconnectivity among participants, who are usually skilled musicians, by allowing them to contribute with music to multiple-user composition sessions, manipulate and shape their's and other players' music, and take part in a collective creation. Interaction in such networks is usually centralized and sequential as participants submit their pre-composed tracks to a central hub and manipulate their peers' material off-line.

This Field Map of Networked Music Systems, is centered in a criteria based on how computer mediation facilitates Musical Interconnection. Therefore, some of the Systems included in the same category might have different approaches in regard to its architecture, functionality or other musical aspects, such as, adequacy for performance or composition practices.

For example the FMOL System (Jordà, S. and Aguilar, T., 1998) and the Webdrum System (Burk, P., 2000b) are both included in the "*Construction Kit Approach*" (Weinberg, G., 2005), even though they are totally different systems in terms of the music creation process.



FMOL is an asynchronous software oriented towards iterative off-line composition (in the iterative database repository usage of this system real-time processing only occurs in the off-line client side), while Webdrum is a Synchronous system oriented towards online jamming and performance (real-time collaboration on the web).

In fact, there are many possible approaches when proposing a systematic field classification and its result will always depend on what information is intended for the reader to extract from this categorization.

One interesting proposal for a high level categorization of Networked Music Systems was suggested by Scot Gresham-Lancaster during an extensive exchange of e-mails between the lecturers of the 1<sup>st</sup> Networked Music Workshop held during ICMC 2005 in Barcelona (Ross Bencina, Jason Freeman, Álvaro Barbosa and Scot Gresham-Lancaster).

“(...) there are two distinct streams here. One that is regarding network music that is intended for a general and possibly non-musician user, and then techniques for practicing musicians to extend the boundaries of performance. (...) they need to be addressed within their separate contexts.” (From an e-mail by Scot G-L June, 17th of 2005)

This proposal contemplates only two categories of systems: (1) Systems and techniques for practicing musicians; (2) Systems for a general and possibly non-musician users. Even though each category covers a very broad spectrum of different features in terms of musical and social interaction practices, this would be a perfectly suitable format of information for a high level user which intends to learn about the best system to use according to his musical skills.

In 2003 the author of this dissertation published a survey of Networked Music Systems in response to Leonardo Music Journal's call for articles entitled "*Groove, Pit and Wave -- Recording, Transmission and Music*".

The Survey entitled "*Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation*" (Barbosa, A., 2003) intends to provide a systematic Classification Space of existing systems, that should be regarded as a starting point for anyone interested to be introduced to this area as a developer, researcher or user.

## 2.4 Networked Music Systems Overview

The Survey presented in this dissertation is by no means absolute, since much of the work done in this emerging field is most likely to be unpublished. Many Artists and Experimental Creators are not concerned with scientific dissemination of their work or research and therefore much information is not available to the public domain.

Nonetheless, the examples presented and discussed in this session, can be regarded as representative of different classes of systems with different specific characteristics related with the requirements of collaborative music practice.

From an attentive analysis of developments in this area during the period of the doctorate work leading to this dissertation (from 2001 to 2005), and using as a reference the classification criteria used in CSCW<sup>13</sup>, the following categories are proposed:

**Co-Located Musical Networks** – Used in organized events for groups of performers who interact in real-time, in the same physical location, on a set of music instruments (or Virtual Music Instruments<sup>14</sup>) with the possibility of sonic interdependency provided by a fast local computer network.

**Music Composition Support Systems** – Used to assist more traditional forms of music composition and production, both for composition oriented towards a written music support or music production based on multi-track and non-linear recording processes. It enhances traditional collaboration paradigms by allowing geographical displacement and asynchronous collaboration.

**Remote Music Performance Systems** – Used in organized events for groups of multiple remote performers/users, displaced in space, improvising and interacting synchronously on a set of music instruments (or Virtual Music Instruments). In this case the sonic interdependency is affected by network latency. A Tele-presence scenario (remote unilateral participation) is a particular case of this set of applications.

---

<sup>13</sup> CSCW is the Research field outside a musical context which deals with computer mediated collaboration in a general way. It is discussed in section 1 of this chapter.

<sup>14</sup> The concept of Virtual Musical Instrument is discussed in chapter 3

**Shared Sonic Environments** – Class of applications that explores the distributed and shared nature of the internet. It is not oriented towards a time limited event scenario. It is more suitable for synchronous improvisation since it addresses the internet community in general providing simple and effective interaction paradigms for collective sonic expression. It does not require previous musical knowledge from the participants, and therefore often results in experimental sonic pieces.

These categories considered in function of the CSCW environmental facets (Synchronous and Asynchronous for the Time Dimension; Remote and Co-located for the Space Dimension), result in graphical representation analogous to Tom Rodden’s Classification Space (Rodden, T., 1991), but representing Networked Music Systems.

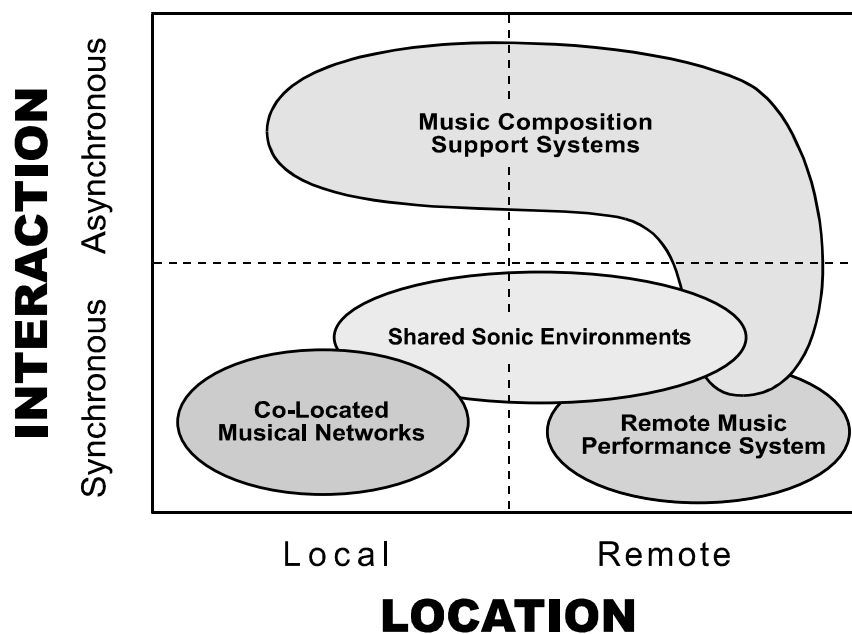


Figure 6. A Classification Space Networked Music Systems

It should be noticed that these are by no means closed categories, and some of following applications could belong to different classes if we consider a less wide-ranging classification criteria.

### 2.4.1 Co-Located Musical Networks

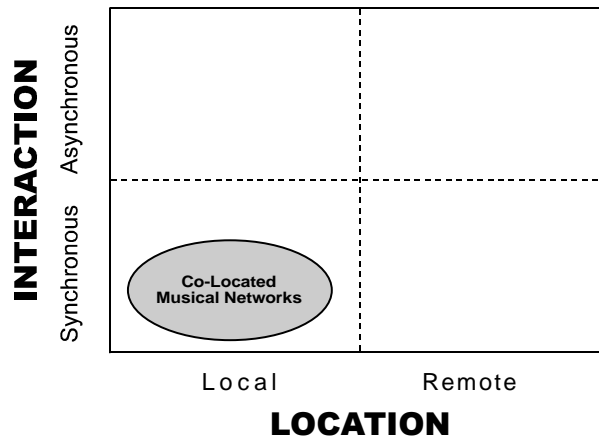


Figure 7. Co-Located Musical Networks in Networked Music Classification Space

This is the class of systems in which the Music Creation Paradigm introduced by the League of Automatic Music Composers and the Hub<sup>15</sup> falls.

However, the research and development conducted by Tod Machover with the HyperInstruments Group at the MIT Media Lab conveys some of the presently most representative examples of Co-Located Musical Networks. Gil Weinberg's work in this Research Group resulted in several notable examples of these type of systems, among which are the **Fireflies** (Weinberg, G., Lakner, T. and Jay, J., 2000), the **Squeezables** (Weinberg, G. and Gan, S.-L., 2001), the **Beatbugs** (Weinberg, G., Aimi, R. and Jennings, K., 2002), **Drum Network**<sup>16</sup> and the **ReacTable** (Jordà, S. and others, 2005)

These types of systems are often bound together with this idea of a Multi-user Musical Instrument. In the history of western music there are very few cases when an instrument was designed to be played by more than one person simultaneously (a piano is often played by four hands even though the instrument was not designed for this purpose) (Jordà, S., 2005b).

<sup>15</sup> The League of Automatic Music Composers and The Hub were presented in section 3.1 of this chapter

<sup>16</sup> The Drum Network provides players with a collaborative playing experience where participants can manipulate, share, and shape each others' music in real time. The drums in the network serve as controllers, sensing hitting and pressure that is then sent via a central system to other players. The drums also serve as speakers by using an attached actuator, which provides acoustic and tactile feedback.

With the emergence of local high speed communication computer networks together with sensor technology, a wide range of possibilities was opened in this field. The Media Lab Instruments mentioned before result from this technological development, as well as many other experimental instruments like the **Jam-O-World** Multi-player Musical Controller (Blaine, T. and Forlines, C., 2002).

However it should be clear that an instrument designed to be played by several performers simultaneously (a Multi-user Musical Instrument), is not necessarily equivalent to a system able to create several instances of the same instrument allowing different users to play together.

It is clear though, that live performance system, which allow players to influence, share, and shape each other's music in real-time or synchronously, are based on a Multi-user Musical Instrument.

In this case, a high degree of interdependence between performances is expected in order to achieve virtuous results, and therefore real-time communication requirements are a critical point. This is the main reason why this approach has been constrained to Local Area Networks, where communication latency allows real-time immediate connections.

## 2.4.2 Music Composition Support System

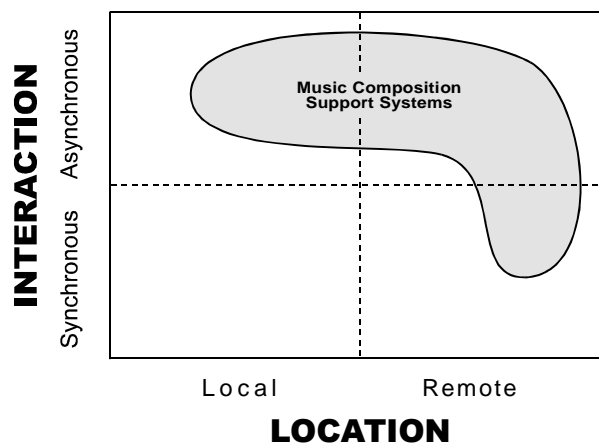


Figure 8. Music Composition Support Systems in Networked Music Classification Space

The primary function that emerged from the use of internet technology in the musical context was to provide mechanisms that assist the composition of music pieces by means of network communication.

Composing music by two or more authors is a process that traditionally can be accomplished in different ways.

Conventionally, to compose a musical piece, a composer individually conceives the music and registers his ideas with symbolic musical notation – a score – (usually in standard western notation). In order to cooperate with other composers in a co-authored piece, it is necessary to exchange ideas amongst contributors by a bilateral analysis of everyone's scores which is only possible if everyone is familiar with the adopted notation.

It is therefore not surprising that one of the first systems, based on the idea of using the internet for enhancing the traditional joint composition in music, goes back to the early 1990's with the Craig R. Latta's **NetJam** (Latta, C., 1991) from Berkley University. This system allowed a community of users to collaborate producing music in an asynchronous way by automatically exchanging MIDI files through e-mail.

A considerable improvement in the efficiency of the traditional symbolic composition process was achieved by introducing asynchronous collaboration combined with geographical displacement capabilities.

#### 2.4.2.1 On-line Music Recording Studios

Another possible approach to compose music emerged in recent years with the advent of recording and editing technologies. The idea of using a Recording Studio as a composition tool became increasingly successful especially in Popular Music.

In traditional composition it is also normal that a composer is assisted by a music instrument when he is experimenting with ideas that will lead to a final result registered in a symbolic support. However, in a recording studio session it is possible for one or more musicians to record their instrumental performances (synchronous or asynchronously), resulting in raw sound material registered acoustically that can be manipulated to a very large extent in order to create a complete musical piece.

This method to create music is highly successful with less trained musicians and composers since it reduces the gap between having an idea and achieving a result, and therefore provides the possibility to react, transform and improvise faster and more efficiently.

This topic has been studied from an Musicological approach, by Paulo Ferreira Lopes, António Sousa Dias and Daniela Coimbra (Ferreira-Lopes, P., Dias, A. and Coimbra, D., 2005) and

these two forms of composing music, in fact, correspond to High Level Models proposed by these authors.

The Endogenous creative trajectory correspond to the traditional composition approach mentioned first at section 2.4.2 and the Exogenous creative trajectory corresponds to compositional approach often practiced in recording studio environments.

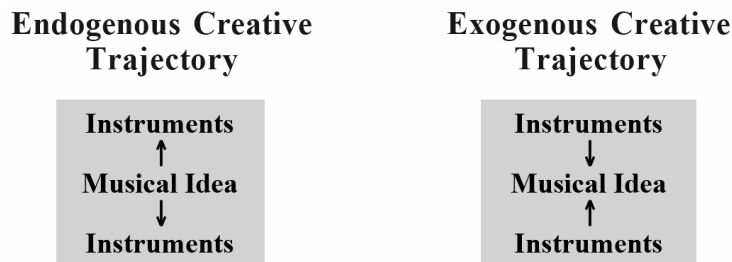


Figure 9. Draft Model of Endogenous and Exogenous Creative Trajectory

For the majority of internet users interested in creating music this process is inevitably more engaging and from this observation a new class of internet applications emerged for music creation aiming to materialize the idea of an on-line Recording Studio.

To take advantage of internet's global communication possibilities in this context, new systems came up based on the idea of collaboration between geographical displaced users in one common project developed in a virtual studio environment.

As a concept these are distributed systems, however there also exists a centralized server that is part of the system and that manages the organization of users into multiple session groups.

Typically the interface layer resembles typical multi-track software, like Digidesign's Protools, Steinberg's Cubase or Nuendo and allows the users to lay down tracks of MIDI and digital audio either in a synchronous or asynchronous mode collaborating with other users that have access to the session.

A pioneer system which followed this approach was the **ResRocket Surfer**. It was a successfully freely distributed application, released in 1994 (Moller, M. and others, 1994), with a reasonable amount of users forming a community of musicians that actually created music cooperatively over the Internet. The system allowed performers and listeners to organize into multiple groups called Virtual Studios, and to lay down tracks of MIDI in an overall composition.

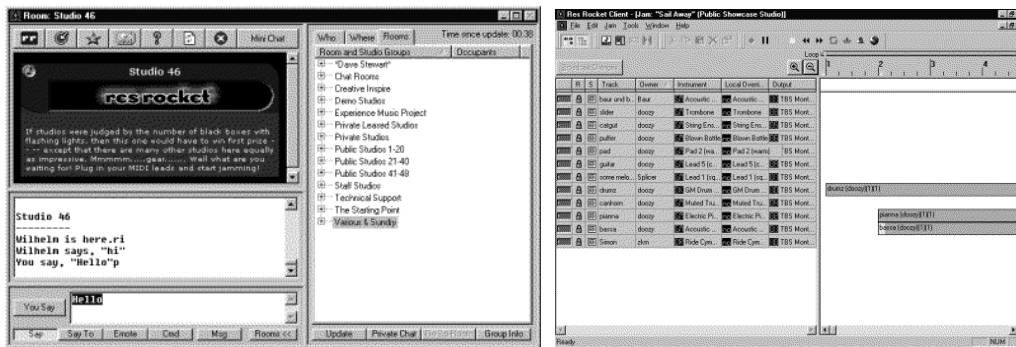


Figure 10. Screen Shots of the ResRocket Software showing a structured list of ongoing sessions and a multi-track project view with individual tracks recorded by different users

It somehow resembled a Chat application where the user could browse to a database of on-line users and join or leave sessions any time, as long as they have the proper permissions defined by the creators of the session. The ResRocket Software could work either in a synchronous or asynchronous mode.

The company Rocket Networks that developed this software, simultaneously introduced the Rocket Power Audio Software that was based on a Centralized Network Model aiming professional recording that supported digital audio and MIDI, which worked in an asynchronous way due to latency over the internet.

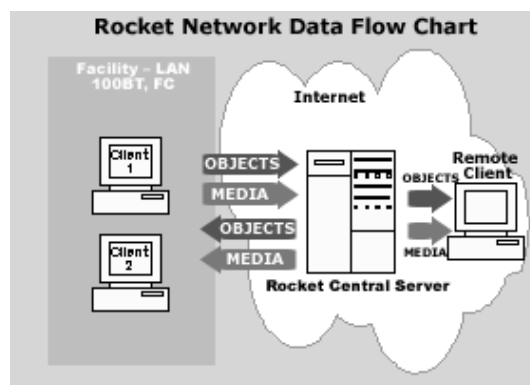


Figure 11. Rocket Power Audio system Topology

Rocket Power allowed the creation of Virtual Work Places and Synchronization is provided by central server, and was becoming the industry standard for the support of long distance collaboration on digital non editing recording software packages, being currently supported by: Digidesign's Protools, Emagic's Logic Audio and Steinberg's Cubase VST. Despite the relative



success in the early days of Internet global usage, the Rocket Networks Company unexpectedly ceased activities in 2000.

In the following year a similar system entitled **TONOS-TC8** was introduced as a commercial Application. Software (TONOS Company, 2001).



Figure 12. TONOS-TC8 Interface

TONOS TC-8 followed the same principles of the basic ResRocket application, but it was restricted to 8 tracks of audio and MIDI recording, allowing users to access a centralized account with 40Mbytes of Hard-Disk Space. Similarly to Rocket Networks this company suddenly ceased activity in 2004.

Even though these systems did not make it as viable commercial products, the basic principles behind their implementation are still very attractive to Popular Music Recording Professionals. Therefore, it is not surprising that in 2005 two new systems based on this paradigm were released.

In April 2005, during the *Frankfurt Musikmesse* (International Trade Fair for Musical Hardware and Software) the German based company **DigitalMusician.Net** (DMN)<sup>17</sup>, presented a software prototype that offers similar features to those of the previously mentioned systems, but using mp3 (128 Bit) compression for digital audio transmission and a video-conferencing system for visual feed-back between users.

---

<sup>17</sup> DMN is available from <http://www.digitalmusician.net/>



Figure 13. DigitalMusician.Net User Interface

Likewise in June 2005 the American Company **eJaming** released their online jamming Studio for MIDI Instruments (Nelson, T., 2005). The system also incorporates a session manager entitled *Lobby* and allows users to join sessions using an interface called *Stage*.

The most innovative aspect of the eJaming Software is the fact that it is the first commercial application to use the **Latency Adaptive Tempo** and **Individual Delayed Feed-Back**<sup>18</sup> concept to diminish the disrupting effect of network delay in performance synchronisation. In the Company's Software Description release is stated:

“(...) Thanks to our patented eJamming™ technology, everybody hears what everybody's playing at each location - in sync, in real time, or in as close to real time as the laws of physics allow.

How do we do it? eJamming™ algorithms delay the sounding of your instrument until you receive data from your fellow eJammers. So from the time you hit your keyboard, strum a guitar string or strike a drum skin, the time it takes to hear that note and those of the other players on your stage varies from 15mS (milliseconds) within a city, 25-40mS within a 1500

---

<sup>18</sup> The Latency Adaptive Tempo and Individual Delayed Feedback concepts are extensively covered in Chapter 4 of This Dissertation and were originally published by the author in May 2005 at the NIME 2005 Conference (Barbosa, A., Cardoso, J. and Geiger, G., 2005)

mile jam and 40-50mS cross country. eJamming™ has even connected musicians from NYC to London at 49mS (...)"

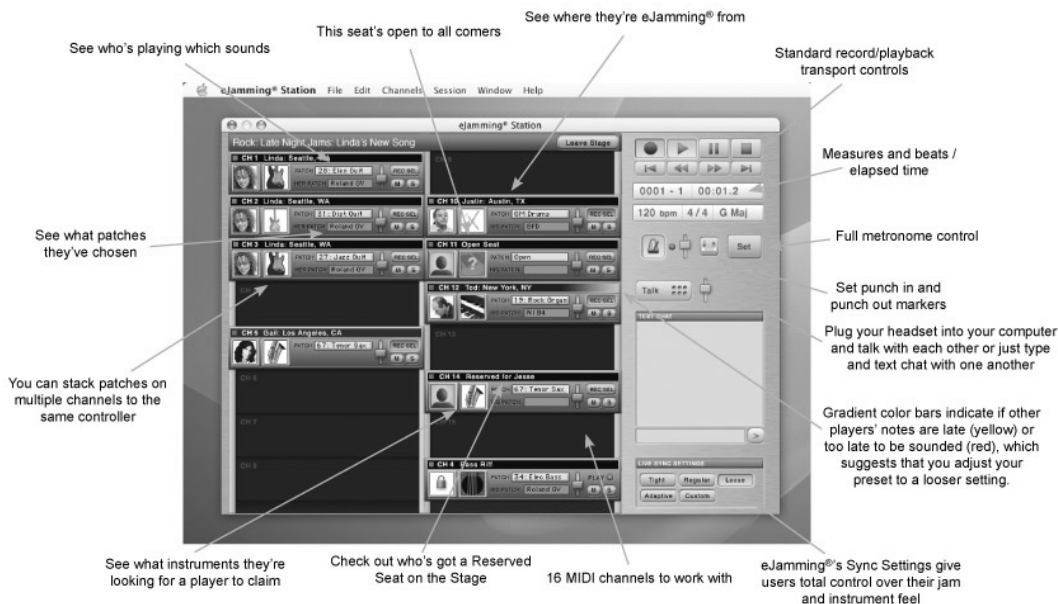


Figure 14. eJamming Stage interface

All these systems are suited for a Music Studio Composition Process, since they have a common option of Multi-Track Recording incorporated with long distance audio communication.

In this experimental composition process composers are often required to develop some sort of written notation to automate the edition and sequences (effects, samples, loops, etc). This practice has the disadvantage of being less universal, since either it is a personal technique developed by experience in studio technology, or a proprietary form of notation imposed by hardware and software manufacturers.

Attending to this requirement in September 2005 another commercial product was released. The **VSTunnel**<sup>19</sup> Software is a VST Plugging (compatible with any audio production software that supports VSTs) used like an insert effect in a sequencer's master out channel, which allows an audio connection over the Internet to other VSTunnel enabled clients.

<sup>19</sup> VSTunnel is available from <http://www.vstunnel.com/>



Figure 15. The VST-Tunnel Plugging Interface

#### 2.4.2.2 Experimental Collective Composition Systems

All previous systems mentioned in the category of Music Composition Support Systems were designed to address a very specific target audience, practicing musicians and music producers which have preceding knowledge in working with Digital Non-linear Multi-Track recording software. However, with the Internet also came a sense of democratization in information access, which empowered, both regular and expert users to create and participate in experimental music composition practices through systems specifically designed for this purpose.

The earliest examples of On-Line Composition Environment, designed for collective music creation by non-practicing musicians came up in the late 90s.

**FMOL** (Faus Music On-Line) is a pioneer Software, developed in 1997 as the result of a commission by the Catalan theatre group La Fura dels Baus to the Experimental Artist Sergi Jordà.

This Software has been developed over different versions until the present day and it can be thought of as a complete system that can be approached from different perspectives. On one hand, as a standalone Electronic Music Instrument it is a powerful paradigm providing unique sonic results that are somehow related with one of the basic initial requirements of the software, which stated:

“(...) it should be a light and fast piece of software in order to be able to run in any inexpensive computer without any additional controllers other than the keyboard and the mouse”. (FMOL initial Requirements)

The sound quality is 22 KHz and the transformation algorithms are (most of the times) simplified versions of high sophisticated sound synthesis filters in order to run fast and provide an immediate response to the users giving them a better feel of playability. This approach introduced sound artifacts to the FMOL sonority, contributing to its unique aesthetical quality. Furthermore the FMOL’s Bamboo interface is also a key element as to what concerns its rhythmical and melodic progressions, which are also unique in this instrument.

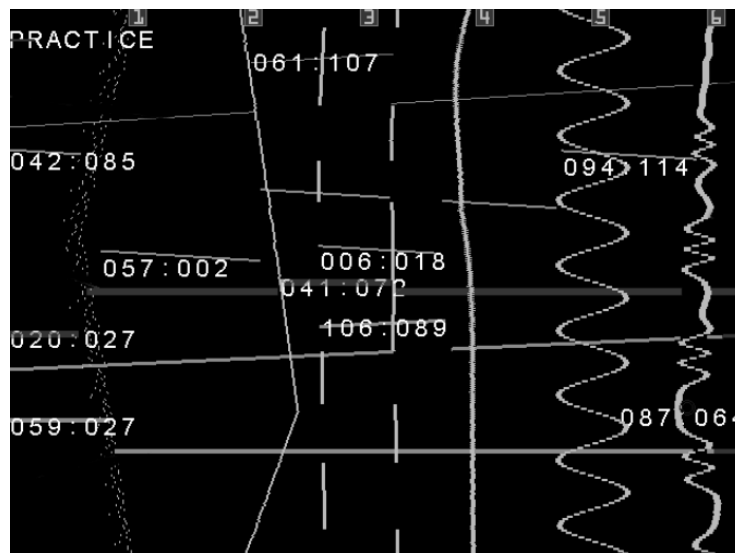


Figure 16. The FMOL Bamboo Interface

Bamboo was designed bearing in mind that its control can be fully mastered, visually resembling a rectangular web, where the horizontal lines correspond to the sound generators, and the vertical lines to the processors. It behaves like a guitar or a harp, as its strings can be plucked or fretted with the mouse, and it also behaves like a multi-channel oscilloscope, since every vertical string continuously draws the sound it is generating.

In Golo Föllmer’s on-line essay “Soft Music” (Föllmer, G., 2001), Sergi Jordà refers to the musical ideas that led him through the process of programming FMOL:

“(...) I keep changing them while I write them and they always surprise me, and the more they surprise me the more I like them. That’s why I like FMOL. It created a musical style that I didn’t know before I started the

program (...)". (Sergi Jordà from a Video Interview in Crossfade, December 2000)

As a live performance instrument FMOL has been used in events by the FMOL Trio in a setup based on two FMOLs, a saxophone player and in many cases other invited musicians. The FMOL experience as a performance and improvisation instrument both in interdependent communication with other traditional instruments or with other FMOL instruments has resulted in refined and exciting electronic music pieces (Feller, R., 2002).

Another requirement of La Fura dels Baus' commission to Sergi Jordà was that the system would allow collective participation of internet users in composition of pieces with the FMOL software which would later be included in La Fura's play F@ust 3.0 and in fragments of the multimedia opera Don Quijote en Barcelona, which premiered at the Gran Teatre del Liceu of Barcelona in October 2000. The original system was built following a client server model, allowing composers using the FMOL client software to log into a central web based server, in order to download pieces stored in a song tree-structure database.

For the collective interaction model a "vertical" approach was preferred instead of a more typical sequential approach which would consist of pasting small fragments one after the other.

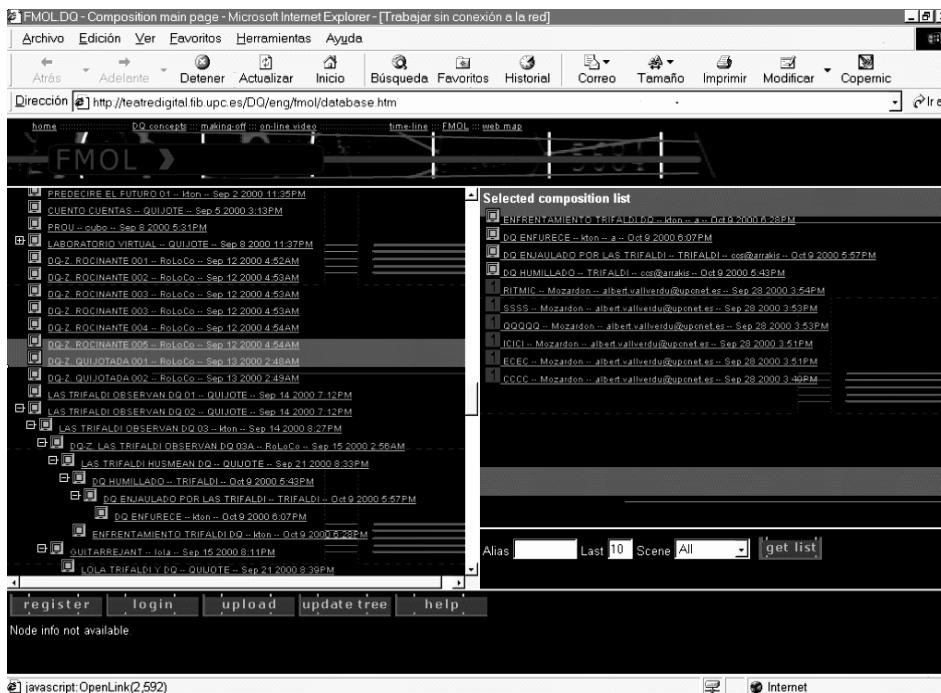


Figure 17. Screen Shots of the FMOL Software showing the web based tree-structured data based with multiple generation pieces composed by different users

In FMOL's collaborative approach it is possible to transform or modify pieces composed by other users, given that the server database structure keeps track of all generations in each composition providing an asynchronous paradigm for co-authored iterative music pieces. Given that FMOL is an accessible system to the regular internet user, since it does not require any special hardware or any special music creation experience, it is perfectly tailored for a internet collective music creation approach (Jordà, S., 1999).

During the two periods in which this project has been on-line, January/March 1998 and September/October 2000, several hundred of composers participated in the active creation of parts for the musical scores of two important plays by la Fura dels Baus, and a collective Audio-CD has been released (Jordà, S., 1998).

Also in 1997, another early and representative system was released in the Internet. William Duckworth's **Cathedral** conceived from scratch to work over the World Wide Web (WWW), even though in 1997 there were fewer than a million sites registered on WWW (Duckworth, W., 1999). The first version of the Cathedral Site included streaming Audio, Video, Animation, Images and Texts. The goal was to create an imaginative, ongoing artistic experience by blurring the distinction separating the composers, performers and audience, and inviting everyone visiting the site to be a creative participant.

The components of this interactive paradigm were the Web site, PitchWeb, and an Internet Band. The Web site features a variety of interactive musical, artistic, and Text-Based experiences; the PitchWeb allows listeners to participate actively and creatively; and the Cathedral Band, which gives periodic live performances and offers listeners focused moments in which to come together and play music in community on-line (Duckworth, W., 2005). The system has been maintained and technologically updated until this day<sup>20</sup>.

One other possible scenario in the context of collective composition of music by communities of users is the case that existing communities with focused interests in Audio and Music end up stimulating mechanisms of interaction to enhance their communication paradigm leading to compositional environments.

---

<sup>20</sup> The present version of the Cathedral in 2005 is heavily based on Macromedia Flash technology. It is available from <http://www.monroestreet.com/Cathedral/>

This is the case of the **Free-Sound Project**<sup>21</sup>, created in 2004 by Bram Jhong at the Music Technology Group of the Pompeu Fabra University in Barcelona, with the purpose of serving as groundwork for the International Computer Music Conference 2005, which was dedicated to the Free Sound Theme. The project is a collaborative database of Creative Commons<sup>22</sup> licensed sounds, with several mechanisms of sound Surfing, Downloading and Uploading. It became rapidly successful leading to further experimentation in terms of the available features on the web site, leading to a particular collaborative mechanism, which became a collective sonic composition tool, the **Remix! Tree**.

The basic idea is very similar to the previously mentioned iterative database model from FMOL. When the users add a sample which is a remix of another sample, it will appear in a tree structure. Remixed samples appear as branches in the tree.

---

<sup>21</sup> The Fee-Sound Project is available from <http://freesound.iaa.upf.edu/>

<sup>22</sup> Creative Commons is a nonprofit organization offering flexible copyright for creative work. (<http://creativecommons.org/>)



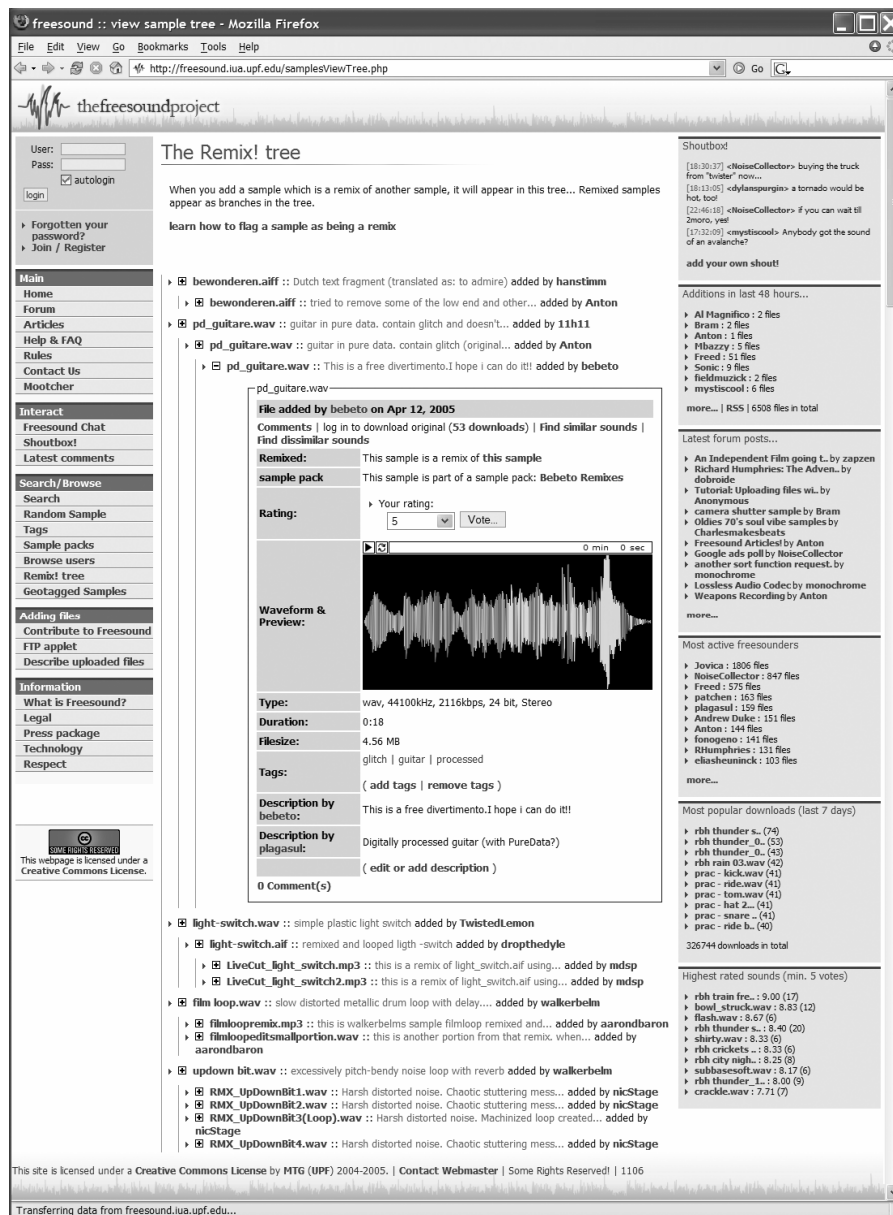


Figure 18. The Free Sound Project Remix! Tree Interface

These systems provide effective enhancements in the process of music production. Yet, they are mostly oriented towards composition perspective, leaving little space for more experimental forms of performative Arts, and thus constraining the potential of what the Internet can offer as a medium for artistic expression in itself.

### 2.4.3 Remote Music Performance Systems

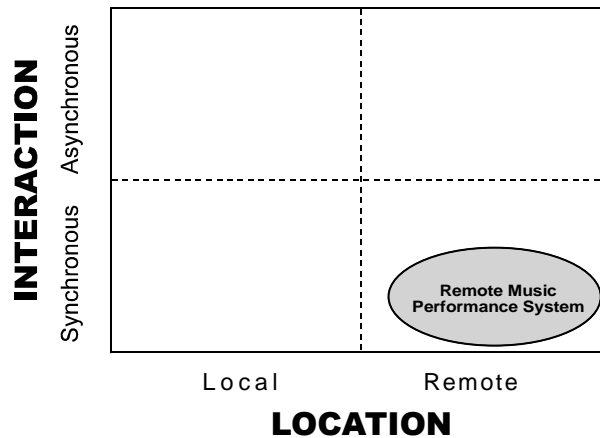


Figure 19. Remote Music Performance Systems in the Networked Music Classification Space

Many examples can be found in which the Internet's potential is explored in order to provide a connection between physical spaces geographically apart.

Nonetheless, there are crucial differences in terms of the complexity of a working system that serves a Broadcast or Unicast scenario in which it is only intended to provide **one-way Tele-Presence**, or a Multicast<sup>23</sup> communication setup that links two, or more, **Collaborative Performative Spaces**.

#### 2.4.3.1 Tele-Presence Systems

The idea of having the presence of one or more remote performers from anywhere in the world in events taking place in physical spaces, facing live audiences, during fixed periods of time is an exciting one.

Of course considerations must be made in terms of when to present public events that occur simultaneously in different places at a global level. If for example a concert would be presented publicly at the East Coast of the United States during the afternoon, it is unlikely that it would be possible to associate this event with another public presentation taking place in Europe, since it would occur in the middle of the night.

---

<sup>23</sup> The concepts of Unicast, Broadcast and Multicast are discussed in chapter 3

Such an event is also very complex in terms of bringing together all the required logistics, since it deals with sensitive, and many times experimental, technology that must comply and work together in distinct sites, separated by long distances. One other concern is network bandwidth, which might be an impediment for a continuous data deliver at the client side.

Different approaches were taken to address this issue in remote performances projects realized over the last few years.

One approach has been to use **cutting edge communication technology**, like high speed and broadband networks combined with streaming technology.

Over the last few years, leading research in this field was conducted by Jeremy R. Cooperstock at the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT)<sup>24</sup> from McGill University in Montreal, Canada and by Chris Chafe at the Center for Computer Research in Music and Acoustics (CCRMA) SoundWIRE<sup>25</sup> Group from Stanford University in San Francisco, USA.

Following the release of 1998 Audio Engineering Society (AES) whitepaper on "Networking Audio and Music Using **Internet2** and Next-Generation Internet Capabilities" (Bargar, R. and others, 1998)<sup>26</sup>, in collaboration with McGill University the **first landmark in Broadband Internet Audio Streaming** was realized in September 26, 1999. A musical performance at McGill University in Montreal, was transmitted over the Internet to a live audience at New York University, during the **107th AES Convention** (Xu, A. and others, 2000).

What made this event distinctive was the audience's experience of uninterrupted, intermediate quality, multi-channel audio (AC-3). In order to achieve this result, a custom system was

---

<sup>24</sup> The CIRMMT Group: <http://www.music.mcgill.ca/cirmmt/>

<sup>25</sup> The CCRMA SoundWIRE Group: <http://ccrma.stanford.edu/groups/soundwire/>

<sup>26</sup> Internet 2 (<http://www.internet2.org>) is a consortium being led by 200 universities working in partnership with industry and government to develop and deploy advanced network applications and technologies, accelerating the creation of tomorrow's Internet. Internet2 is recreating the partnership among academia, industry and government that fostered today's Internet in its infancy.

developed employing both TCP and UDP protocols<sup>27</sup>, and providing its own buffering and retransmission algorithms.

Following this experiment at the **109th AES convention** in Los Angeles, the world's first transcontinental studio experiment was demonstrated on Saturday, September 23, 2000. This session was entitled "**The Recording Studio that Spanned a Continent**" (Cooperstock, J. and Spackman, S., 2001).

The McGill Jazz Orchestra performed in a university concert hall in Montreal while the recording engineers mixing the 12 channels of audio during the performance were not in a control room at the back of the hall, but rather across the continent in a theater at the University of Southern California in Los Angeles, mixing for a live audience.

This was the first time that live audio of this quality was transmitted over an Internet2 network with the resolution of 96 kHz/24 bits linear-PCM. Once the 12 channels of audio were mixed into six 96/24 outputs in a digital console in the USC theater, the six signals were converted to analog by 96/24 D/A converters before being sent to the theater's 6.1-monitoring system (Woszczyk, W. and others, 2005).

A different approach to performance in a live public event is to incorporate **low-cost public domain technology**. Although it is less appealing if one is trying to realize a traditional music experiment, it has been the basis for most expressive experimental artistic performative events over the internet.

Different styles of music, instruments and technical setups have been tried like the Telemusic and the **Piano Master Classes** by John Young and Randall Packer (Young, J. and Fujinaga, I., 1999) (Young, J., 2001) or the New York University's **Cassandra Project** (Ghezzeo, D. and others, 1996).

Both these projects feature situations of remote performance in live events where the demarcation of physical and virtual space, on-line and local proximity converges and blurs into a shared participatory experience, raising questions not only regarding sonic aspects but also about what should be the actual remote performer's visual representation on site.

---

<sup>27</sup> The TCP and UDP protocols are briefly discussed in chapter 4

A more complex scenario than a unilateral Tele-presence, discussed so far is, a performative bilateral collaboration between two simultaneous events.

### 2.4.3.2 Collaborative Performance Systems

An example of a possible (and typical) setup of a performative event which separates practicing musicians over a long distance connected through the Internet is presented in the following Diagram.

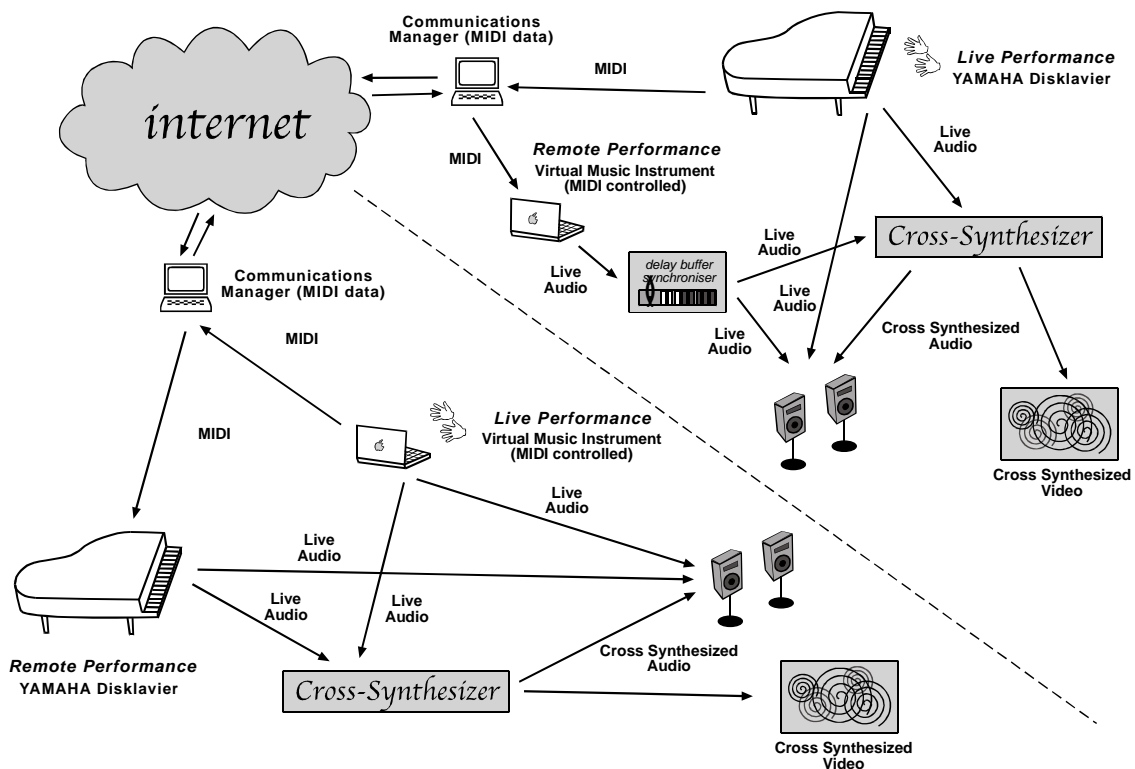


Figure 20. Diagram for a remote collaborative musical performance for a pianist (using Yamaha Disklavier Pianos) and live electronics performed on a Laptop.

This project implied a collaborative performance between *Fundamenta Nuova* Theatre in Venice, Italy and *Auditório Ilídio Pinho* at the Portuguese Catholic University in Porto, Portugal.

The Performance required high-end Musical Instruments (Yamaha Disklavier, MIDI controllable, Piano), video synthesis and a Virtual Music Instrument Setup. It shows that in this class of Remote Music performance Systems the Logistics, hardware requirements and the overall complexity of getting everything to work properly is not of a simplistic nature.

The 1987 “**Clockworktower Concert**” by **The Hub**<sup>28</sup> is unquestionably one of the earlier, if not the earliest, examples of a geographically displaced collaborative music performance with computer technology (Brown, C. and Bischoff, J., 2005), and throughout the decade similar experiments were carried out, but unfortunately few were documented or disseminated in the Music Technology community.

Some known examples of the Collaborative Performance approach are<sup>29</sup>:

**Eve Schooler: Distributed Music: A Foray into Net Performance (Sept. 1993):** Synchronized three real-time streams from different hosts; delays in the order of 200 ms difficult for performers to be listeners

**Paul Hoffert: CyberSoiree (Feb. 1996):** ATM-based technology for audio and video streaming of a four-way jazz performance. Delay>0.5s delay but musicians earned to compensate through extensive practice.

**Dimitri Konstantas: Distributed Musical Rehearsal Studio (May 1996):** ATM based distributed rehearsal with conductor at different location from musicians. 80 ms one-way delay for audio-video synch; echo resulted in "extreme confusion" (Konstantas, D. and others, 1997) (Konstantas, D., 1998)

**Seiji Ozawa: Opening Ceremony Nagano Winter Olympics (1998):** Conduct choruses on 5 continents: singers in Sydney, New York, Beijing, Berlin, False Bay. Time lag adjustor used to eliminate satellite delay

Still in 1998, during an interview for the Computer Music Journal with the **Sensorband ensemble** , Zbigniew Karkowski refers to the group’s extensive experience in collaborative performances over the internet (using ISDN connections) which they called Network Concerts conveying the fundamental artistic issues raised by this set-up:

“Another artistic aspect of ISDN concerts is the idea of control. Very often, composers use computers to achieve greater control. We have found, after playing several concerts like this, that we could never control the output

---

<sup>28</sup> This concert is briefly referenced in section 2.3.1 of this chapter

<sup>29</sup> From Jeremy Cooperstock’s notes on Internet Audio Landmarks – ANET Summit Presentation, 2004.

100 percent. Aspects like the delay become unknown variables, which is interesting (...)" (Bongers, B., 1998)

In the same interview Atau Tanaka adds that:

"As artists, our first instinct is not to make technical improvements to the system, but rather, to manipulate the technology in a creative manner. The technical limitations become characteristics of the composition. Doing this allows us not to be so worried about transmission delay, rather, to be concerned about the general notion of distance (...)" (Bongers, B., 1998)

The SensorBand concerts were based on synchronous collaboration in a peer to peer model between two performers. Yet, other experimental systems focused on the idea of having several synchronous performances as close as possible to a real-time situation.

This is the case of the 1998 **TransMIDI** system (Gang, D. and others, 1997), implemented using the Transis multicast group communication layer for CSCW applications (Amir, Y. and others, 1992). This system allows musical performers (and listeners) who wish to play together to organize into multiple session groups.

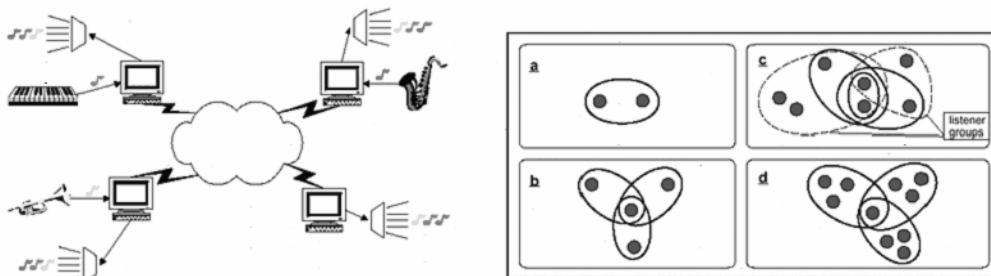


Figure 21. Diagrams from the TransMIDI System showing possible group topologies

Users interact in a synchronous mode (close to real-time) over the network, and may dynamically join or leave a session group. Players contribute to the session by playing on their MIDI controllers, using General MIDI protocol, and it is possible to have different topologies including the formation of hybrid groups of participants and cores with one or more leaders, also permitting access to listener groups.

Some of the most recent Collaborative Music Performance experiments, many of which are in the context of music education, took advantage of broadband technology and videoconference technology<sup>30</sup>:

**Internet2 Initiative: World's First Remote Barbershop Quartet (Nov. 2000):** Multi-location quartet; each of the 4 singers in different cities, conductor in 5<sup>th</sup>. Network delay variances prevented singers from hearing each other or conductor.

**Internet2 Initiative: Music Video Recording via Internet2 (Nov. 2000):** Multi-location music video recording session using real-time streaming video. Musicians simultaneously connected via timing tracks to a mixing board.

**Chris Chafe: QoS Enabled Audio Teleportation (Nov. 2000):** CD quality sound (750 kbps) of 2 separated musicians in Dallas streamed to Stanford. Musicians played "together" in same space (Stanford) but delay was severe.

**Zukerman "Playing Together" Sessions (Dec. 2000):** From New York-Ottawa, Pinchas Zukerman teaches violin classes to McGill University students, in Montreal, using broadband connectivity from the National Research Council in Ottawa to the McGill University in Montreal.

**John Wawrzyniek: Network Musical Performance (May 2001):** Gestural coding (e.g. MIDI) used to manage data for distributed musical performance. Musicians at Berkeley and CalTech, playing on MIDI keyboards; local feedback only.

**Zukerman Music Master Classes (Feb. 2002):** Again from New York-Ottawa, Pinchas Zukerman teaches violin classes to McGill University students, in Montreal from the National Research Council in Ottawa to the McGill University in Montreal, but this time using broadband CA\*net3 (Canadian fiber-optic network), capable of transmission rates of up to 40-gigabits per second. This allowed the use of SDI video (High resolution Digital Video) and multi-channel 96 kHz/24 bit audio and display on 50" plasma screen (near life-size). Improvements in immersive perception were remarkable.

---

<sup>30</sup> Partly from (Woszczyk, W. and others, 2005) and from Jeremy Cooperstock's notes on Internet Audio Landmarks – ANET Summit Presentation, 2004.



**McGill-Stanford Jam Sessions (Jun. 2002):** The UltraVideoconferencing technology developed at McGill University was used in a cross-continental jam session between Musicians at McGill University and Stanford University. The event featured full-screen bidirectional video and multi-channel audio in what was the first demonstration of its kind over IP networks.



Figure 22. Musicians at McGill University (Dan Levitin – sax and Ives Levesque – trombone) Jamming with remote Musicians at Stanford University projected on Screen (Alexander Carôt – Bass and Estabin Wilson – sax).

Furthermore, if we consider all the experimental Musical Practice that has been carried out in the last decade using video and audio conference technology it is most certain that this is the approach to Networked Music that has the largest repertoire of music performed.

Innumerable references to this approach can be found in **Sot Gresham-Lancaster** article “*Video Conferencing Software as a Performance Medium*” published in 2005 at the Networked Performance Blog from the Turbulence.Org Web Site<sup>31</sup>.

A similar scenario can be found in communities of experimental Artists and Engineers which develop custom made experimental music software, frequently adapted for internet Joint Performance. A representative example is Sergi Jordà’s **FMOL Peer-to-Peer** multi-user development based on the original software mentioned in the previous section of this chapter.

---

<sup>31</sup> The article can be found at [http://www.turbulence.org/blog/archives/2005\\_04.html](http://www.turbulence.org/blog/archives/2005_04.html) (accessed 2005/10/08)

This enhancement implemented in 2001 consists in providing MIDI connectivity over the Internet between two Instances of the software running in remote machines. All the sound computation is done by each peer.

The first FMOL Networked concert took place in October 2001, during the Networkshop festival in Dresden, Germany, between the location of the festival and the city of Barcelona, Spain. Attained delays were in the range of 100 ms using a conventional 56 kb modem connection, and according to Sergi Jordà a very good feeling of playability was achieved with this setup.

This condition of immunity to network delay in FMOL music is related to the nature of its free and improvisatory Musical structure. The sound sequencing technique used in this system, based on low frequency oscillators (LFOs) excitation of sound generators, creates rhythmical and melodic progressions which, to some extent, support flexible reaction times and short lacks of synchronicity from the performing partners.

Another recent example of custom made software musical instrument designed from scratch to be used as a collaborative tool is the **PeerSynth** software, developed in 2003 by Jörg Stelkens (Stelkens, J., 2003). It is a unique software piece, which establishes a Peer-to-Peer network through the TCP/IP Protocol<sup>32</sup>.

---

<sup>32</sup> Peer-to-Peer Networks and Communication Protocols are discussed in chapter 3

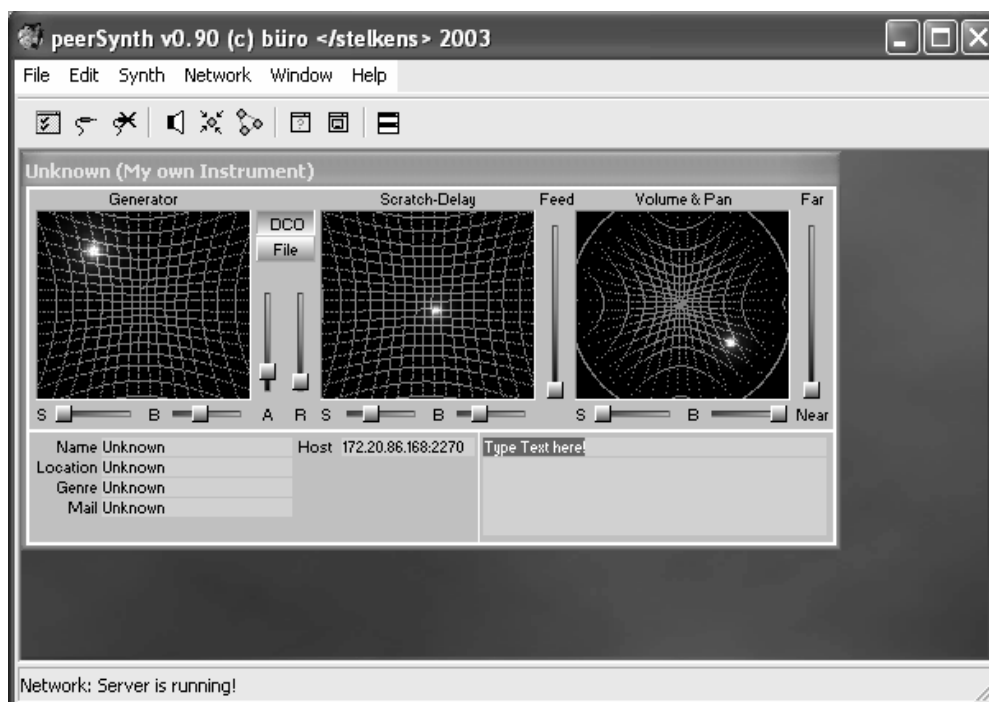


Figure 23. Screen-Shot from the Peersynth network synthesizer.

The software supports multiple users displaced over the Internet, measuring the latency between each active connection and dynamically lowering the sound volume of each user's contribution in the incoming soundscape, proportionally to the amount of delay measured in his connection.

Stelkens followed a real world metaphor where, in fact, the sound volume of a sound source decreases with the distance to the receiver, which also implies increasing acoustical communication latency.

Yet another example of the same approach is Georg Hajdu's **Quintet.net** project (Hajdu, G., 2003), in which five performers can collaborate through the net by means of a custom developed Max-MSP client front-end. In the development of this system, special care was taken to accommodate different musical approaches ranging from free improvisation to the performance of compositions with fixed notation (Hajdu, G., 2004).

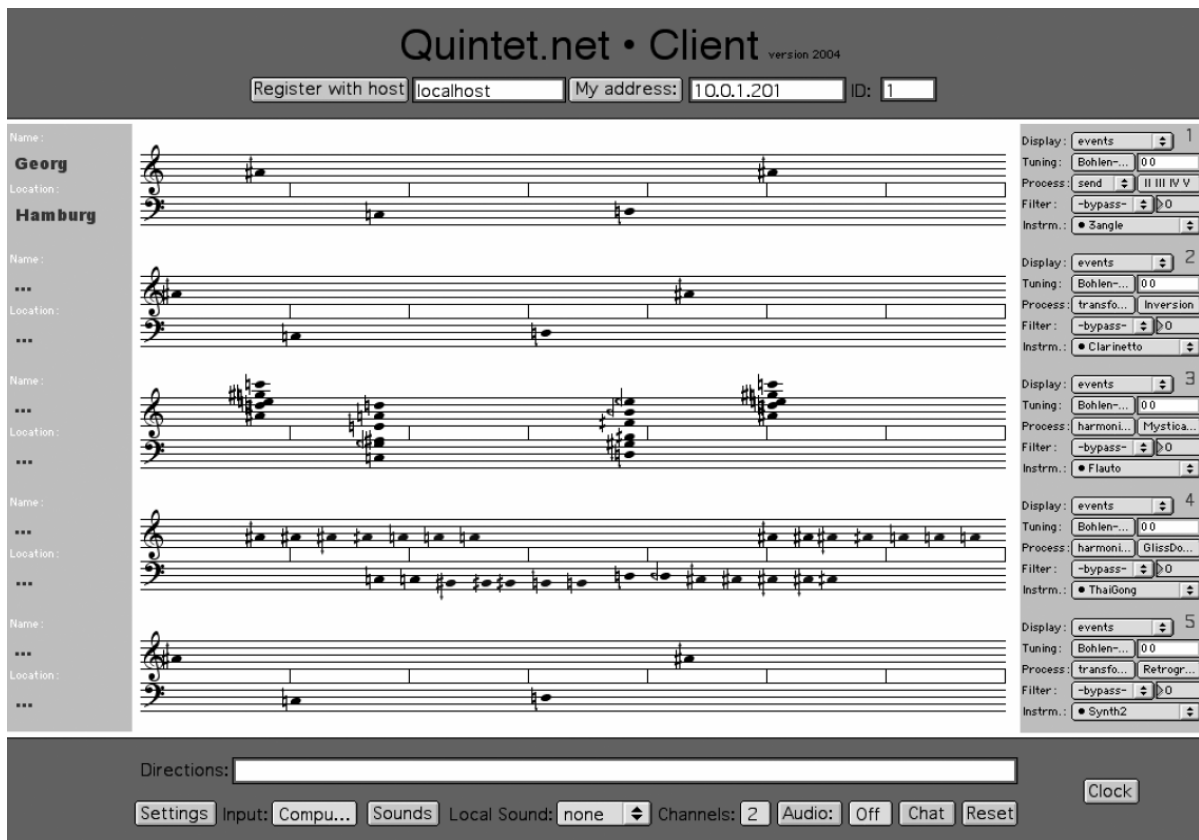


Figure 24. Screen-Shot from the Qintet.net client interface.

Systems designed for real-time collaboration and jamming on the Internet can follow diverse technical approaches to minimize the disrupting effects of latency. Nonetheless, it is also possible to intervene at the musical level to overcome this problem. This is what was achieved by the project **NINJAM**, written by Cockos Incorporated and Brennan Underwood in 2005<sup>33</sup>.

The system is based on a Novel Intervallic Network Jamming Architecture for Music. It departs from a basic principle of forcing latency to become a multiple of the musical tempo measures, allowing users to play together synchronously even though they won't be playing in the same tempo interval as other player.

<sup>33</sup> NINJAM is available from: <http://www.ninjam.com/>

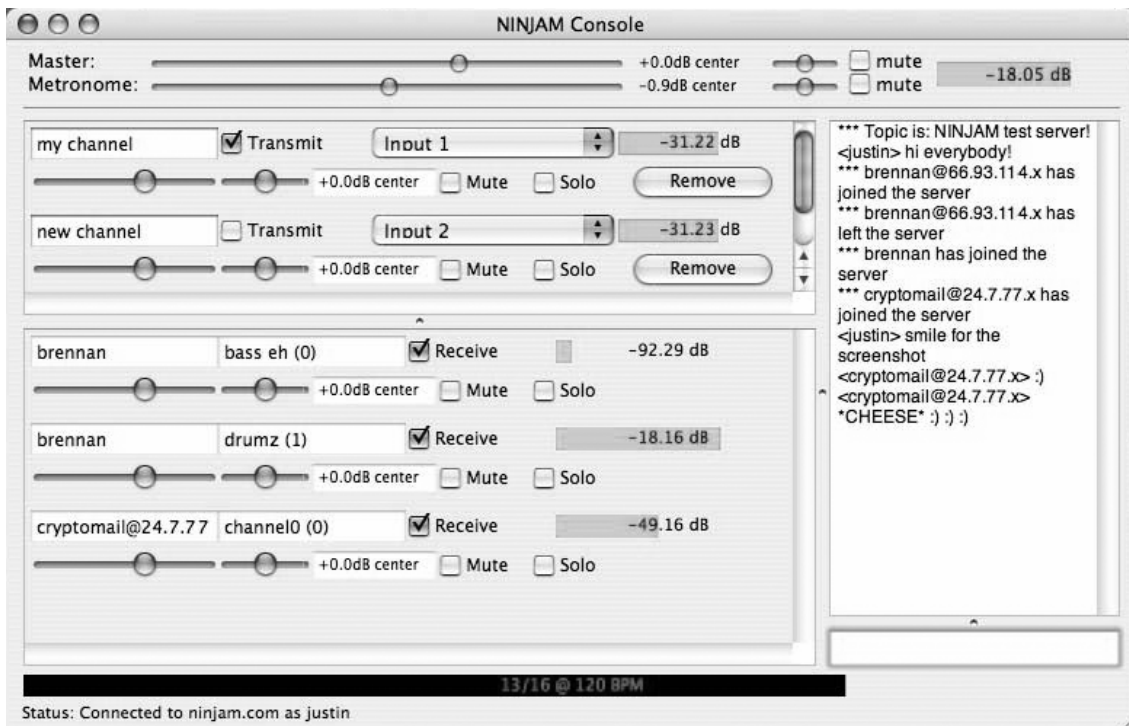


Figure 25. NINJAM client interface

NINJAM uses compressed audio which allows it to work with any instrument or combination of instruments. It streams compressed audio to a NINJAM server, which can then stream it to other performer in a jam session.

The Remote Music Performance approach to the development of Networked Music Systems mainly uses the Internet as a communication media that provides links between performative spaces in an event driven perspective and performed by a well specified group of users.

A different way to move towards Networked Music over Internet is by exploring its shared nature by the means of distributed On-line shared spaces suitable for collective sonic creation by anonymous, possibly non-musicians, on-line users. This approach leads more towards improvisation paradigms as discussed in the following section.

## 2.4.4 Shared Sonic Environments

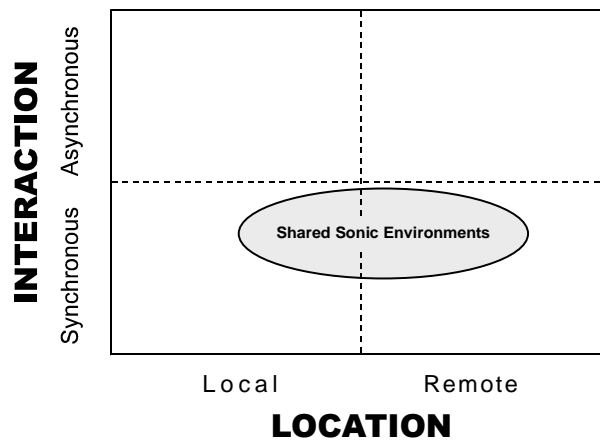


Figure 26. Shared Sonic Environments in the Networked Music Classification Space

Shared Sonic Environments is pulled out from the concept of Shared Virtual Environments (SVEs) discussed in section 2.1.4 of this chapter and embraces several distinctive features:

- The focus is on synchronous collaboration between on-line users and usually more than one user is connected at any given moment (locally or remotely)
- It is based on a public shared space that is openly available to the online community and therefore it must use the most disseminated and open technology on the Internet.
- People can be found on-line improvising in collective music pieces, given that everyone should be able to choose either to participate as a performer or simply as a member of the internet based audience
- No requirements are demanded from a regular user in terms of previous knowledge of musical practice
- Each user is normally able to express himself by somehow manipulating or transforming a sound or a musical structure
- It is suitable for a spontaneous improvisation approach to sonic creation.
- Due to the permanent availability of these systems it supports events which are unlimited in time.

Before presenting some of the work done in this area we find it useful to discuss two of the inherit concepts from outlining Shared Sonic Spaces: On-line Improvisation and the Time scales of a permanent event.

#### 2.4.4.1 On-Line Improvisation

Given that, regardless of the latency, Internet communication media is tolerably suitable for synchronous collaboration, thus Remote Music Performance Systems, discussed in previous section, fulfill the requirements for an event oriented towards remote performance. Furthermore, the physical setup requirements for a remote performance should be the same for **musical interpretation or improvisation**.

Yet, even though the setup requirements are physically the same in both cases, there are major conceptual differences.

In the context presented in this dissertation, by **musical interpretation** it is meant the process of playing a predetermined sequence of events in a musical instrument providing some sort of synchronism with other musicians or audiovisual events. In a musical interpretation there is a great deal of space left for individual expression and even for an improvisation experience, however the events performed by the musicians are driven by a prearranged sonic choreography to a very large extent.

In **musical improvisation** musicians are not coupled in such a systematic approach, and much more space is available for spontaneity, free expression and continuous development of elaborate interactive relationships between the participants. This process is many times referred to as a Jam session. One can also think of improvisation as the process that results from composition coming together with interpretation, in the sense that when improvising the musician is creating musical structures with a sense of awareness similar to the composition process, even though he is doing it in real-time, reacting to an outside stimulus, like when interpreting music.

Technically there are no strict boundaries between interpretation and improvisation. Instead there is a continuum in which the musician drifts according to the context of the performance.

In the book “Composing Interactive Music”, Todd Winkler (Winkler, T., 1998) refers to different performance models: The Conductor Model (used in Symphony Orchestras conducted by one single individual); The Chamber Music Model (used in String Quartets where

conductor's role change amongst the performers during the development of the piece); The Improvisation Model (used in Jazz Combos, based on a free interpretation of written music); The Free Improvisation Model (also used in Jazz in a totally spontaneous interaction paradigm).

We can immediately identify a strong resemblance with some approaches followed in the Remote Music Performance Systems, described in the previous section of this chapter, and realize that systems like TransMIDI were designed for performance, but they also support improvisation.

On the other hand, one could also think that in the age of synthesizers and software generated music, the role of human performance based on written music could be questioned, especially in systems that are highly difficult to control since a synthesizer, or a computer can reproduce any symbolic notation flawlessly even with high degrees of detail to preserve the complexity of the music pieces. In this perspective what remains for human performers is to introduce their own expressiveness and spontaneity, which are the basis for improvisation.

It in this perspective, the emergent new class of applications proposed in this dissertation defined as **Shared Sonic Environments** complies with the idea of a suitable paradigm for music improvisation.

#### 2.4.4.2 The Time-Scales of a Permanent Event

In the book *Microsounds* (Roads, C., 2001) published in 2001 by MIT Press, Curtis Roads introduces a detailed taxonomy of timescales from a music theory perspective.

In Curtis Roads proposal, timescales in music are decreasingly the Infinite timescale, the Supra timescale, the Macro timescale, the Meso timescale, the Sound Object timescale, the Micro timescale, the Sample timescale, the Subsample timescale and the Infinitesimal timescale.

Most musical creations driven by the concept of an event are situated in the Macro timescale defined by Roads as “The time scale of overall music architecture or form, measured in minutes or hours, or in extreme cases, days”. However, one could question where an ongoing musical piece, permanently available for hybrid communities of creators and listeners belongs.

Realistically this scenario should fit in the Supra timescale, which Roads defined as “A timescale beyond that of an individual composition and extending into months, years, decades, and centuries”, since the Infinite timescale is in reality a mathematical abstraction and it is beyond the time life of the present cultural and technological state of development.



Some recent Artistic proposals approached the concept of a musical event that is unlimited in time.

In 1999 Antoine Schmitt<sup>34</sup> created the Infinite CD for Unlimited Music, the first infinite CD to be published and distributed. This CD produced by Epplay, Schmitt and Icono, once inserted in a computer generates music infinitely, always different and always similar, without any images or any form of interaction.

Another essential reference is on-line piece entitled “Eternal Network Music” by Chris Brown and John Bischoff<sup>35</sup>, where several flexible music pieces permanently go on since February 28, 2003.

In the same way that the Internet’s essence is to provide permanent connectivity, a Shared Sonic Space event is also permanent and public, since it is continuously available to the public both via the Internet providing a permanent choice for the users to be in either the performer’s or the spectator’s role.

#### 2.4.4.3 System Implementations

A very Early example of a Shared Sonic Environment System implementation, which was significantly inspiring for the early research work leading to this dissertation, is Atau Tanaka’s **MP3Q** piece on the web (Tanaka, A., 2000), classified by the author as a shared on-line sound space. The web application streams multiple channels of mp3 audio from different servers and users can concurrently manipulate these mp3 sources by actuating over graphical representation of the systems current behavior via a sort of 3D Cube.

An extremely significant couple of developments that led to a number of applications centered on the idea of a shared Sonic Environment were two very specific technologies envisioned by Phil Burk: An audio Software Synthesis Application Programming Interface (API) for Java entitled **JSyn** (Burk, P., 1998), allowing multi-platform client sound synthesis in web-browsers,

---

<sup>34</sup> More information about Antoine Schmitt’s infinite CD can be found at the web site <http://www.infiniteCD.org/>.

<sup>35</sup> These pieces are available from the Web Site: <http://crossfade.walkerart.org/>

and the **TransJam** System (Burk, P., 2000a), a Java based server that can be incorporated into applications that allow people to interact synchronously over the Internet.

Based on these Technologies three Shared Sonic Environment Systems were implemented<sup>36</sup>:

- The **WebDrum**, developed by Phil Burk, is a drum box that can be shared by several people over the Internet. Users are not required to have any musical experience and while sharing this software they can chat with each other at the same time as editing drum patterns and listening to the music they create together.
- The **Eternal Network Music**, developed by Chris Brown and John Bischoff, is based on two interactive music pieces, which are part of a historical retrospective of the Hub. In fact, the Hub's work inspired the development of the TransJam server.
- The **Auracle**, developed by Max Neuhaus, Phil Burk, Jason Freeman, C. Ramakrishnan and Kristjan Varnik, is a voice driven, interactive, collaborative instrument. Working from Stuttgart Germany, they have used JSyn and TransJam technology along with Linear Predictive Coding (LPC) speech analysis, neural nets, evolutionary strategies and other techniques to create an engaging sonic environment (Ramakrishnan, C., Freeman, J. and Varnik, K., 2004).

---

<sup>36</sup> Links to this systems are available from the TransJam website: <http://www.transjam.com/>

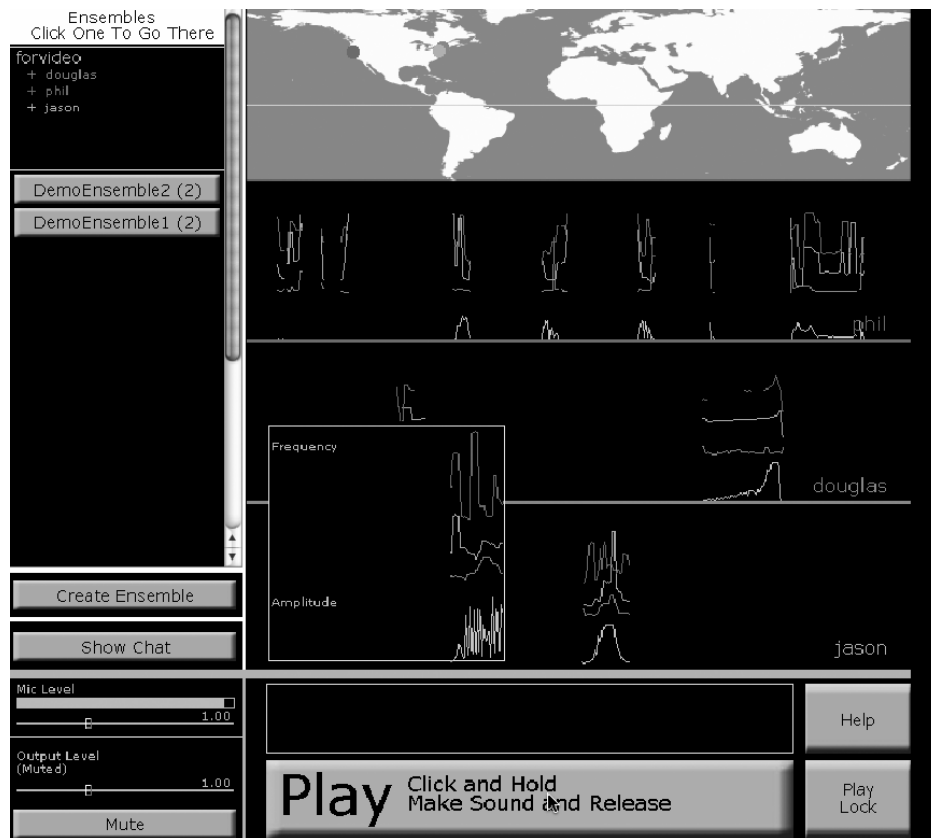


Figure 27. Screenshot from the demonstration video Documentary on the Auracle

An even more extreme example of custom made music software, oriented towards a very specific community of musicians, are live coding environments (programming languages oriented to live music generation), such as Supercollider, or **ChuckK** (Wang, G. and Cook, P., 2003).

The most recent development in the ChuckK framework aims to make this collaborative system oriented towards geographical displacement following the approach of a Shared Sonic Environment. The **Co-Audicle** (Wang, G. and others, 2005) is defined by the authors as a collaborative audio programming space, for collaborative, multi-user interaction based around the ChuckK language.

It operates either in client/server mode or as part of a peer-to-peer network. The different instances conveyed in a Co-Audicle session are represented graphically at the client's interface through engaging metaphors.



Figure 28. Graphical Representation of different computational instances in the Co-Audicle

Shared Sonic Environments is a central concept in this dissertation. In this context most of the experimental research developed between 2001 and 2005 was focused on the development of a Proof-of-Concept system prototype entitled **Public Sound Objects (PSOs)**. The project is a web-based Shared Sonic Environment, which has been an experimental framework to implement and test different approaches and concepts for on-line music communication, discussed in Chapter 3 of this dissertation, such as, the notion of a network-music instrument incorporating latency as a software function, by dynamically adapting its tempo to the communication delay measured in real-time. PSOs first released was in 2002 by Álvaro Barbosa and Martin Kaltenbrunner at the Music Technology Group from the Pompeu Fabra University in Barcelona – Spain (Barbosa, A. and Kaltenbrunner, M., 2002). The System is extensively discussed in Chapter 4 of this dissertation.

## 2.5 Chapter Conclusions

A preliminary review of Computer-Supported Collaborative Systems for music and Sonic Art Creation was released by the author at the MIT Press Leonardo Music Journal 13 in 2003. However the survey presented in this dissertation has been widely revised, expanded and updated.

One should be aware that this chapter is not exhaustive, and due to the experimental nature of this field it is clear that information about many interesting and relevant projects has probably never been published or publicly disseminated in any other form.

Nevertheless, it presents an extensive review and classification of representative projects, aiming to provide valuable references and concepts for future work in this area.

It is possible to infer from this study that most of the projects approaching geographical displacements over the Internet are oriented towards:

- (a) The creation of networks where documents in digital audio or logical formats can be exchanged amongst geographically displaced contributors, in project oriented collaboration paradigm like in a typical Computer Supported Cooperative Work application;
- (b) Providing a channel for tele-presence between performative spaces and therefore enhancing the efficiency of traditional collaborative paradigms for music performance/composition, music education or even for music sharing, by adding long distance collaboration possibilities.

In this context a different approach has emerged, going beyond the enhancements on existing acoustic communication paradigms, and focusing on a diverse breakthrough aspect of Internet collaboration, its shared nature:

- (c) The possibility to create community oriented Shared Virtual Environments, where users can dynamically join and leave a group in a collaborative ongoing sonic performance based on the simple manipulation of sound objects from a soundscape, or even on the creation of musical structures.

Like similar paradigms oriented towards visual or textual communication (MUDs, MOOs, IRC, Active Worlds, etc) tend to lead to new mechanisms of interaction not usually seen in “real life” (Curtis, P., 1992), a similar result can be expected in paradigms oriented to music or sonic arts, suggesting that the sonic outcome of such systems could express interesting new artistic results.

It is clear that this area of sonic creation is quite promising, not only by the fact that it allows the enhancement of known paradigms to make music, but also since it provides a context for stylistic novelty.

The results from this survey led directly to the interest and developments in the Shared Sonic Environment project Public Sound Objects introduced in Chapter 5 of this dissertation.

# Chapter 3

## Networked Music Practice Topologies

In his book “Networking the World”, 1974-2000, Armand Mattelart characterizes communication networks as:

“(...) an eternal promise symbolizing a world that is better because it is united. From road to rail to information highways, the belief has been revived with each Technological generation (...)”

For Mattelart Networks are systems that facilitate the movement of persons, material and symbolic goods, which can have diverse structures (linear, radial, centripetal, rhizomatic) but do not require a bidirectional stream of movement in each channel (Chandler, A. and Neumark, N., 2005).

Similarly to cooperative work mediated by computer technology, Music Practice Networking should be regarded as a paradigm which requires this bidirectional flow of information. In this chapter, general ideas and proposals for generic topologies are presented in the perspective of providing orientation and reference concepts to project development of Computer Mediated Networked Music Practices.

### 3.1 Networked Models for Collaborative Music Practice

The ubiquitous nature of communication in computer networks, firmly manifested in the Internet era, has greatly contributed to a favorable environment in which joint editing systems accomplished exponential acceptance by the on-line community.

Joint editing is the process of developing a Multimedia Document by more than one author, communicating partly or wholly via computer networks.

In this case the main objective is the exchange of partial content which will add up to a final result. However, when introducing the notion of Cooperation in such systems, the idea of Group Communication is also implied. In a Group Communication system users strongly interact during the course of a Multimedia Object creation, in such a way that their work converges to a final result, which would not be the same as the simple sum of their isolated partial contributions, even if they were strictly developed according to the project specifications.

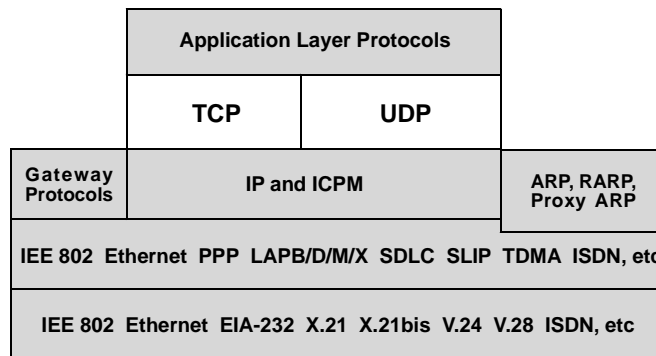
Furthermore, in such processes an isolated analysis of the user's individual work does not convey the same meaning, when integrated in the context of the completed project.

This chapter discusses different communication models and concepts at the structural level, providing a general basis for the implementation of cooperative paradigms.

### 3.1.1 Common Network Protocols, Architectures and Models

For the various users in a distributed multi-user application to share the same virtual space and interact, their host machines must communicate with each other via a network. While there are many different protocols available, two of the most commonly used are the Transmission Control Protocol (TCP) and the User Datagram Protocol (UDP). TCP was initially developed for use in the ARPAnet and later, the Internet, but it is now widely used throughout the world in commercial and academic networks. UDP is simpler protocol and it is used instead of TCP in a number of applications when the full services of TCP are not needed.

The following figure illustrates TCP and UDP operation at the Layered Model and its characteristics, as presented in (Black, U., 2000).



**TCP**

- End-to-end accountability of traffic (ACKs)
- Extensive flow control operations
- Sequencing of all traffic into and out of layer7 applications
- Support of internet port operations

**UDP**

- No end-to-end accountability of traffic
- No flow control operations
- No sequencing of traffic
- Support of internet port operations

Figure 29. Transmission Control Protocol (TCP) and the User Datagram Protocol (UDP)

As we can see, both TCP and UDP reside in the transport layer of a conventional layered model. However, TCP guarantees reliability in end-to-end delivery of messages sent, while UDP gives no guarantees. UDP does not have either flow control (to guarantee the order of the packets) or error recovery measures, serving only as a minimum level service that performs as a multiplexer/demultiplexer to send and receive IP traffic to well identified sources and destination ports.

For these reasons TCP is much slower than UDP, and therefore less suitable for real-time communication. Real-time applications, like multi-user virtual environments, which require the high-speed of UDP can't rely on all message transmissions being successful, or use a hybrid TCP/UDP approach to send slow but reliable messages when necessary.

Furthermore, common problems when using UDP in a multi-user system over the internet, like the **KeyStroke** – Multi-User Cross Media Synthesizer<sup>37</sup>, are related with net routers that

---

<sup>37</sup> The KeyStroke system (<http://www.keyworx.org>) requires fast speed connections, since it enables real-time collaboration between artists from different disciplines exchanging and transforming different kinds of multimedia objects like digital audio, video, static images and text. Players can be working together in the same physical space and connected via a local network on which case UDP is the best option. They



typically come with a default configuration that rejects any traffic that does not have end-to-end reliability of traffic (ACKs). In the case that one of the intervenient computers in a joint performance with a system like KeyStroke, is located at a sub-network behind a router configured this way, it is necessary to use a TCP connection, even though it is slower, since UDP does not provide a reliable transfer function.

### 3.1.1.1 Reliability and Quality of Service

Reliability carries different meanings for different applications. For example, in a typical replicated database setting, reliability means that messages can never be lost, and that they should arrive in the same order at all sites.

In order to guarantee this reliability property, it is acceptable to sacrifice real-time message delivery and some messages may be greatly delayed, and at certain periods message transmission may even be blocked.

While this is perfectly acceptable behavior for a reliable database application, this behavior is intolerable for a reliable video server. For a continuous MPEG video player, reliability means real-time message delivery, at a certain bandwidth. It is acceptable for some messages to be lost, as long as the available bandwidth complies with certain predetermined stochastic assumptions.

In this case introducing database style reliability (reliable message recovery and order) may violate these assumptions, rendering the MPEG decoding algorithm incorrect.

Quality of Service (QoS) reflects the reliability in a specific system. QoS was largely ignored in the initial design of IP when the only requirement was that the data should not be corrupted or get lost. Today with the possibility of transmitting real-time data over large bandwidth IP networks, it is extremely important to be able to characterize and control jitter and end-to-end delay<sup>38</sup> (Hernst, O., Gurle, D. and Petit, J.-P., 1999).

---

can also be connected through the Internet and distributed throughout the world using TCP most of the times.

<sup>38</sup> End-to-end delay corresponds to the bidirectional latency of data transmission between two points. Jitter is a time base error that comes as a consequence of the variable and unpredictable nature of this latency, which creates timing variations in the received data stream.

If real-time transmission of data is not a requirement, the main parameters characterizing QoS in regular a packet network are latency, bandwidth, packet loss and de-sequencing.

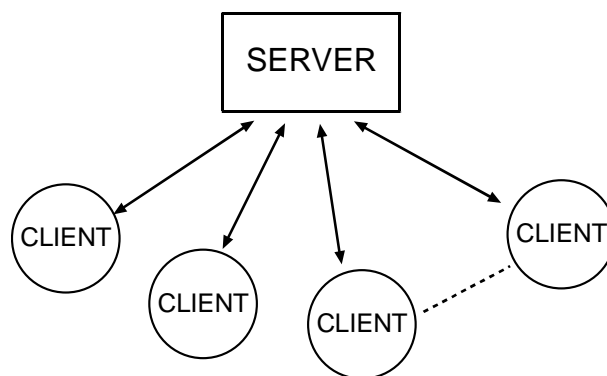
A CSCW application in a musical collaboration context incorporates various activities such as digital audio or symbolic musical representation transmission and management of replicated interaction settings . These activities obviously require different QoS, and yet are part of the same application. Furthermore, CSCW applications in general often need to be fault-tolerant, and need to support smooth reconfiguration when parties join or leave.

### 3.1.1.2 Network Communication Models

When setting up a multi-user collaborative system over a computer network, there are different models for network level data distribution. These models have mostly been developed in order to solve specific problems encountered during the development of Internet. However, they also have profound implications in the way the group will behave together and the results to be achieved.

#### **The Centralized Network Model** (Pulkka, A., 1995)

A client/server, or centralized network model represented in figure 2.3-1, burdens a single host with the task of communicating with each of the clients<sup>39</sup> to determine and report the current state of the system. The server simply maintains the database, while the clients handle computation and rendering.



---

<sup>39</sup> In this context, by client it is not necessarily meant as an application layer controlled by a user, but instead it is often the case that this topology is applied to the interconnection of autonomous computational instances (processes).

Figure 30. Centralized Network Model

This is typically the easiest approach to implement, but it is not scalable. As the number of users increases, the performance between the server and each of the clients decreases.

One way to overcome this scaling problem is to create multiple communicating servers, illustrated in the following figure.

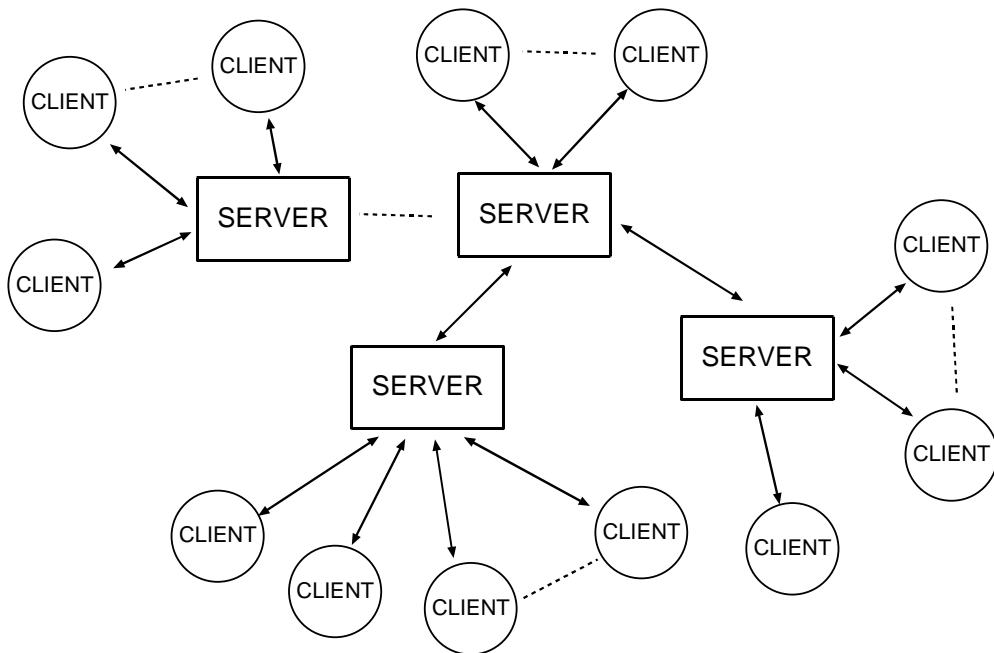


Figure 31. Centralized Network Model with Multiple Communicating Servers

Each client communicates directly with the closest (in terms of network distance) server, which takes care of communicating updates with the other servers, which in turn communicate with each of their clients. This increases the complexity of maintaining a coherent database, but decreases the impact of adding new clients (as long as there are enough servers).

An example of a public domain centralized network model is the Squid Web Proxy Cache<sup>40</sup>, a system which can be configured in a hierarchical fashion or as a mesh.

---

<sup>40</sup> Squid (<http://www.squid-cache.org/>) is a full-featured Web proxy cache, designed to run on UNIX systems and it is a free open-source software. Squid is derived from the ARPA-funded Harvest project developed in Colorado University.

**The Distributed Network Model** (Hernst, O., Gurle, D. and Petit, J.-P., 1999, Obraczka, K., 1998, Pulkka, A., 1995)

A serverless, peer-to-peer, point-to-point, or distributed network model makes no distinction between clients and servers. Each user maintains a local copy of the database as well as handling computation and rendering.

When changes are made to the database, the user must communicate that change to all other users in the system.

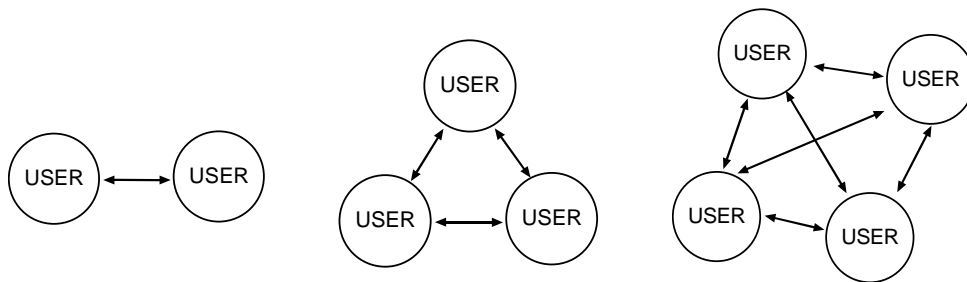


Figure 32. Distributed Network Models

In the simplest case where we only have two users, the process of information transmission is classified as **UNICAST**, since for each update message sent there is only one sender and one receiver.

However, when increasing the number of users the approach for information transmission must be **MULTICAST**, since each time an update message is sent, it must be sent to all the users in the network.

This approach has a scaling problem because the number of messages being sent by each user steadily increases with the number of users, as each user has the responsibility to physically send the messages to all the others.

Techniques can be employed to help reduce the number of messages sent by determining which peers will be interested in any given update message.

### **The Group Communication Paradigm**

Group communication is a powerful paradigm for joint editing of documents. It introduces the notion of group abstraction which is based on the concept that one can consider multiple

connections between the users as a whole and in that sense each user only needs to have one communication channel opened with this group protocol.

In this case the process of information transmission is classified as **BROADCAST**, since rather than sending update messages to each of the other users, each user only sends a single message that is received by all other users in the system.

This paradigm results in fewer total messages being sent over the network, however, broadcast communication has the negative side effect of sending each message to everyone on the network, including those not participating in the virtual environment simulation. This can cause an overwhelming burden on the networks processing abilities.

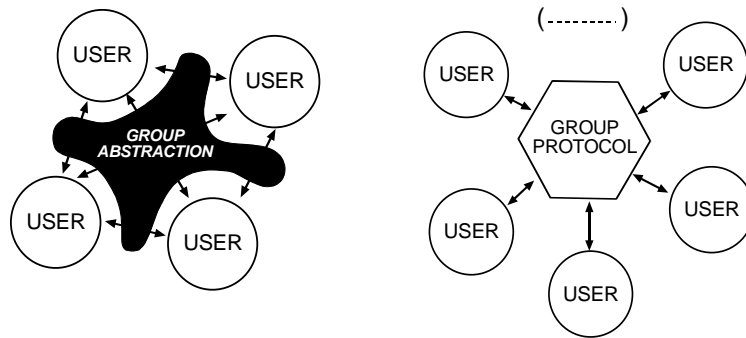


Figure 33. Group Abstraction for Broadcast Transmission

An example of a public domain group protocol is the Spread Wide Area Group Communication System<sup>41</sup>.

### 3.1.1.3 Decentralized Communication Environments

Network decentralization comes with peer-to-peer distributed systems and it was introduced in the internet's original design in the late 60's. The first few hosts on the ARPANET were independent computing sites with equal status and connected together as computing peers and not in a client/server relationship (Oram, A., 2001).

---

<sup>41</sup> Spread (<http://www.spread.org/>) is a toolkit and daemon that provide multicast and group communications support to applications across local and wide area networks. It is designed to make it easy to write groupware, networked multimedia, reliable server, and collaborative work applications.

Over time the internet has become increasingly client/server due mostly to the expansion of the World Wide Web (WWW) that compels to the usage of this model. However, a new generation of applications came up with Napster software and internet started being used much as it was originally designed.

Even though the public attention to the Napster phenomenon was originated by its famous legal problems, the system has probably been the most successful amongst internet users, applying the principle of peer-to-peer data transfer and belonging to a class of applications that take advantage of resources (storage, cycles, content, human presence) that are available at the internet's terminal computers.

The decentralized model allows the development of applications that provide faster and more efficient group communication systems. Yet it is necessary to realize that this model is not applicable in every group paradigm.

In (Oram, A., 2001), page 28, is pointed out that:

“... In fact, peer-to-peer is distinctly bad for many classes of networked applications. Most search engines work best when they can search a central database rather than launch a meta-search of peers. (...) Any system that requires real-time group access or rapid searches through large sets of unique data will benefit from centralization in ways that will be difficult to duplicate in peer-to-peer systems ...” (Oram, A., 2001)

Recent systems have taken this into account and developed applications that work in a mixed model of centralization and decentralization.

In Napster a centralized server maintains a master list of all the song files, but the song themselves are maintained in the clients, with massive redundancy, and file transfers are peer-to-peer.

Centralizing the pointers and decentralizing the content is a powerful paradigm, since it couples the strength of a central database with the power of distributed storage, and it is a very promising proposal for applications in the field of collaborative music creation.

Another approach in distributed computation is in peer-to-peer networks on which the terminal applications are not controlled by users, but instead they are autonomous computational instances. There are many contexts in which this idea provides major results, like in the case of

the SETI@home project<sup>42</sup>. However, in the case of musical and sonic creation it's hard to conceive a scenario where it could be applicable, with the eventual exception of the algorithmic composition field.

### 3.1.2 General-Purpose Models for Music Collaboration

Different network topologies can be applied to different classes of applications. Depending on system requirements the best suitable communication model could be a Centralized Network Model, a Distributed Peer-to-Peer Network Model or a mixed setup. From the classification space proposed by the Author of this dissertation, network communication architectures can be profiled for each class of applications

#### **Co-Located Musical Networks**

Due to the real-time constrains of these systems, and the fact that they are oriented to small groups, the most common network topology is Peer-to-Peer.

#### **Music Composition Support Systems**

These systems could either be implemented with Distributed Asynchronous Paradigms (like e-mail based systems) or Centralized Network Models (systems based on structured multi-user file repositories).

#### **Remote Music Performance Systems**

Due to the fact that these systems are usually based on well-defined groups of users and that even though participants are geographically dispersed communication is required to be as close to real-time as possible, usually Peer-to-Peer topology is employed.

#### **Share Sonic Environments**

In these systems a Shared Virtual Space is required. Therefore it is typical to have some sort of Centralized Instance in order to maintain permanent track of the system status. The Public

---

<sup>42</sup> The SETI@home Project (<http://setiathome.ssl.berkeley.edu/> - Search for Extra Terrestrial Intelligence at home) attracted millions of people long before Napster and it's a system that exploits the enormous amount of idle time in numerous personal computers in a distributed paradigm.

Sound Objects Project (Chapter 5), for instance, is totally based on Client/Server architectures; however, a more general approach would lead to a mixed setup between a Peer-to-Peer connection for content exchange amongst users, and a centralized Instance to keep track of the current system state.

One question that was raised during the Public Sound Objects project development, was whether the acoustic feedback available to the users (the overall acoustic piece streamed back from the server affected by network latency) was accurate enough to give the feeling of playability to the end user.

Part of the problem might be that the individual acoustic feedback for each user's own performance is only being perceived together with the incoming audio stream.

The Centralized Network Model architecture proposed in the following figure addresses local feedback at the client level by the means of local synthesis, and replication of interaction settings amongst users with a Mediator server service.



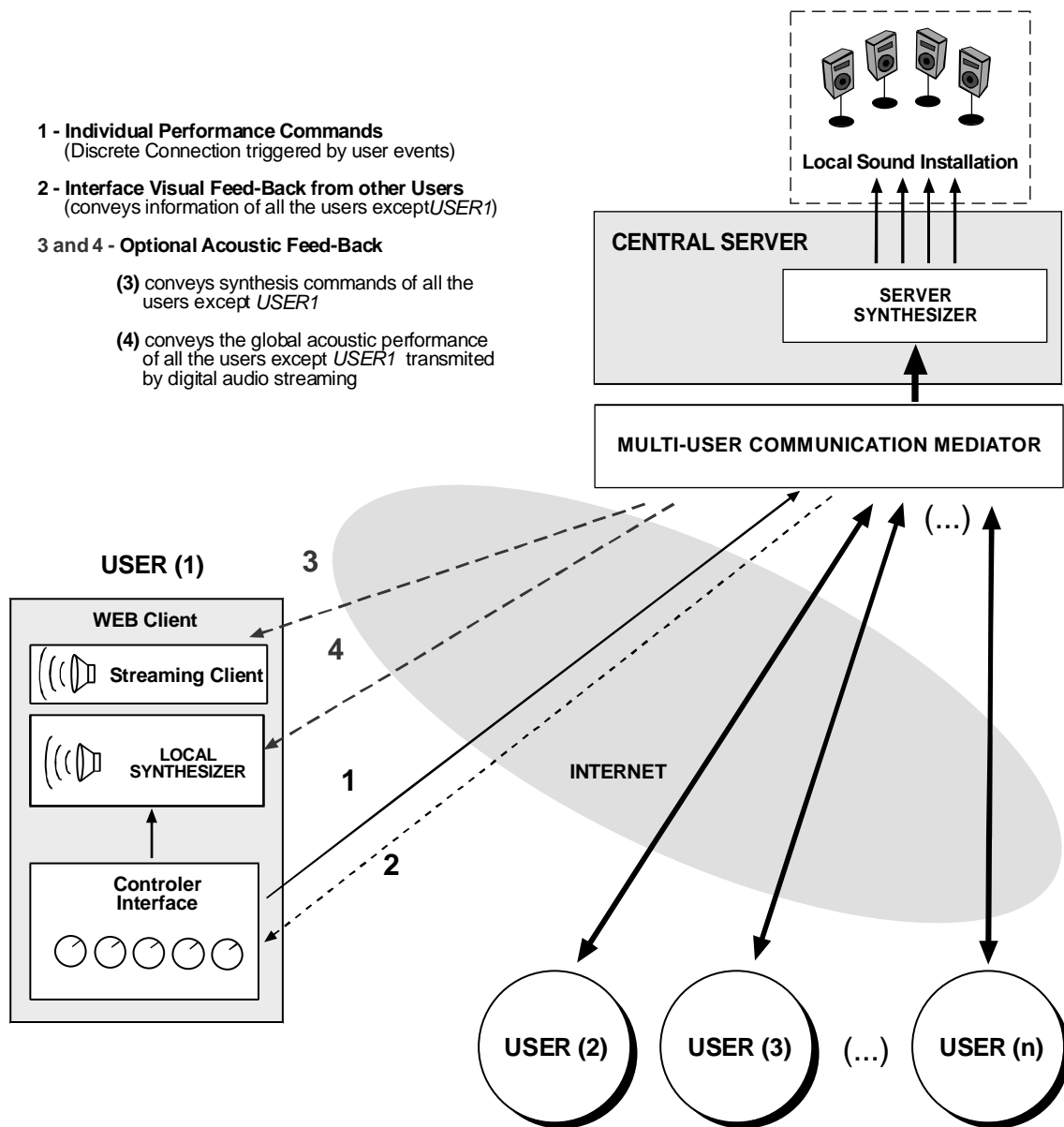


Figure 34. Centralized Shared Sonic Environment Model with Local Feed-Back

Another possible approach for such a system is to use a mixed model with a Peer-to-Peer approach for the exchange of content and a Central Server Instance to keep track of the system settings, presented in the following figure at a higher level diagram.

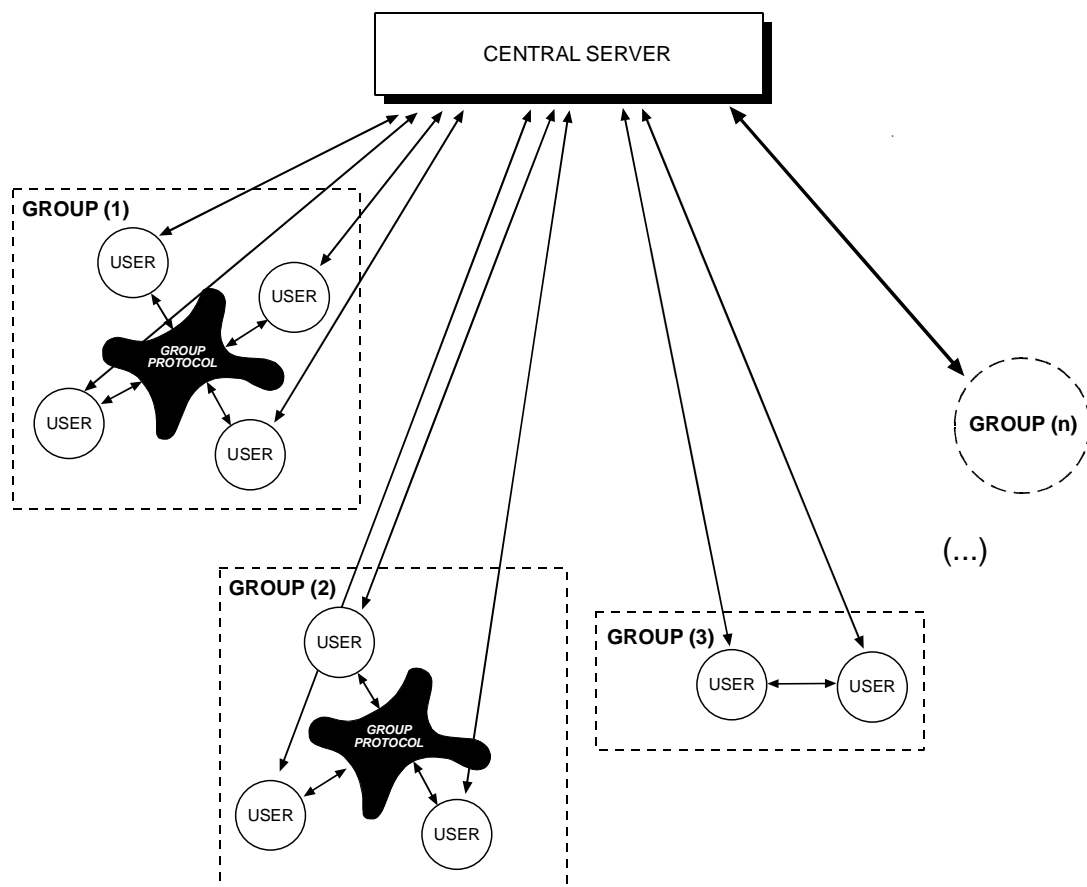


Figure 35. Centralized Sonic Shared Environment Model with Local Feed-Back

The Centralized instance also has an important role in the initial establishment of a user session in order for each individual to choose in which group he would like to participate or which sonic elements he wants to use in his performance.

In these proposals there are still some unanswered questions. Since each user client listens to slightly different versions of the ongoing piece and at the server there is yet another version of the global piece, one could wonder which one is the final produced piece or even if it makes sense to think about one main piece.

Presently existing software applications can work as group protocols with specific application to musical communication (acoustical and Logical).

This is the case of Phil Burk's **TransJam** (Burk, P., 2000a, Burk, P., 2005), a High Level Real-Time Networking Java solution managed by a centralized server.

Transjam has Client and Server Instances which are implemented in Java, and it allows: Client Management (Usernames and logins, Rooms of users); Message Passing (Chat-style text messages, Shared data objects); IP-to-location service (world map); Pre-built GUI components to handle common tasks.

TransJam is the basis for the implementation of systems such as the webdrum or the Auracle (presented in Chapter 2), however it has limitations:

- It has no management of timing and latency
- It is tricky to transmit non-text data
- It has an inefficient transmission protocol
- The GUI components are not very interesting
- It needs permission to run persistent server-side application

Another technology which was extremely well accepted by the Computer Music Community is Matt Wright's **Open Sound Control** (OSC) (Wright, M. and Freed, A., 1997), developed at the Center for New Music and Audio Technologies, U.C. Berkeley<sup>43</sup>.

OSC is a communication protocol by that enables remote connectivity between many synthesis and transformation software used by computer music creators, such as, Pure-Data, Max/MSP or Supercollider. OSC features:

- Binary format for packing “messages”;
- Messages identified by textual paths e.g. /synth1/lfo/frequency/set;
- Messages with multiple parameters, with a number of types, overloading is possible;
- Transport over UDP, but other options are possible (TCP, serial, etc).

As musicians design their instruments they can construct complex connections between dislocated musical systems and easily allow incoming data from other musicians to have control over certain aspects of their software or self-made instrument.

---

<sup>43</sup> OSC is available from: <http://cnmat.cnmat.berkeley.edu/OSC/>.

## 3.2 Towards an Ubiquitous Virtual Music Instruments

Computer Technology extended the notion of a **Multi-User Instrument** and of a **Virtual Music Instrument**. Yet, Networking of musical instruments through digital technology in the Internet era brought the notion of a **Ubiquitous Musical Instrument**, i.e., a Musical Instrument that can have geographically displaced components.

To fully understand an individual interaction model for such an Instrument, it is useful to briefly discuss the interaction models of a performer with a traditional music instrument or a virtual music instrument.

In principle talking about musical performance (or temporal control) implies an instrument and a performer. The individual interaction of one user and a traditional music instrument can be modeled with several levels of detail.

The performer and the instrument form a complex sensorial Feed Back system in which the performer actuates on the instrument, and the changes that occur in the generated sound will influence the performer's actuation in the following moment.

However, the characteristics of the sensorial feed back system between the performer and the instrument go beyond auditory events since the physical and visual interaction also play important roles.

It is also important to notice that this system evolves to different stages, along with the training that is required for a performer to control the instrument on a consequent way.

### 3.2.1 Traditional Music Instruments VS Virtual Music Instruments

The aim of an instrument is that somehow a performer can control it with body gestures in order to produce music. Nevertheless, there are differences between a traditional music instrument and a virtual music instrument:

**Traditional Musical Instruments** – The focus is usually on pitch and dynamics changes. The performer can actuate on the instrument by changing its notes and their amplitude. The general timbre of the instrument is modeled on the physical characteristics of the instrument and if it can be changed, usually it is not in an interactive way during the performance.

**Virtual Music Instruments** – With electronic/digital media developments, especially in computer technology the possibility to control every parameter that modifies sound became possible. Yet, even today there is a tendency to recreate traditional music instruments interaction model which focus on pitch and dynamics.

Transformation of timbre using computers through spectral modeling of sound is probably the greatest potential of what we can call a Virtual Music Instrument.

The Interaction Models for these two classes of Music Instruments also have specificities. The diagram representation of a performer actuating over a Traditional Music Instrument is a simple feed back Loop.

Real time auditory feed back can either be the result of the acoustic transformation that the room applies to the instrument sound, mixed with the direct propagation of the sound from the instrument straight to the performers ears, or in the case of an anechoic chamber or a setup with amplification and headphones only the direct sound that outputs from the instrument.

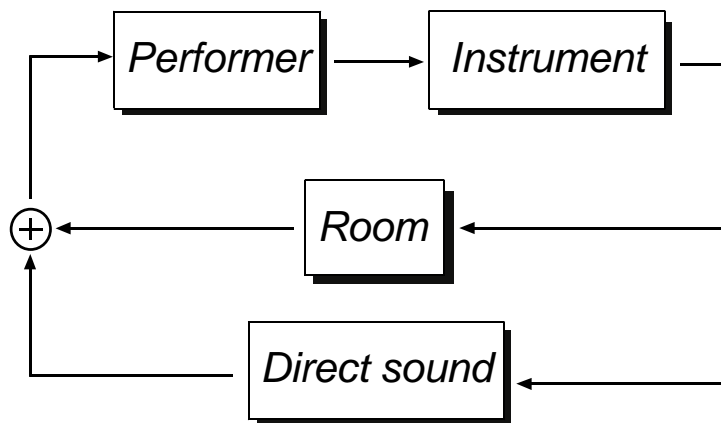


Figure 36. Traditional Music Instrument Interaction Model

Although it is not represented in the diagram, the sensorial Feed Back that the performer receives from the instrument while playing is not only auditory. The physical structure of the instrument also has great influence on the physical gestures of the performer.

A more detailed representation of the sensorial feedback system can be made; however, it will be unique for each instrument since it concerns its physical structure.

The following figure represents a diagram for a model of structural interaction with a violin, presented by Daniel Trueman in his 1999 PhD thesis published by the Department of Music at Princeton University (Trueman, D., 1999).

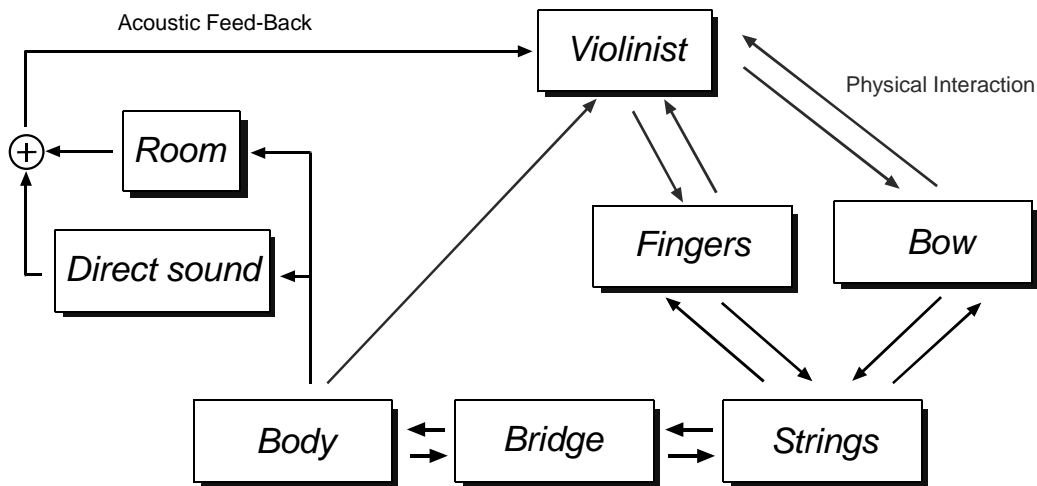


Figure 37. Violin Interaction Model

For Virtual Music Instrument it is possible to define a model for a general system; however, it should be noticed the feedback between the output of the virtual device and the user since it must be in real time so that the player can have the same feel of playability of a traditional music instrument. This is a critical requirement since even with fast computer systems, the processing might be too heavy and the system can be affected by latency over the minimal thresholds required.

A Virtual Music Instrument (VMI), as described by Axel Mulder in (Mulder, A., 1994), aims to provide a way to control parameters of sound synthesis in an expressive and artistically meaningful way. This requires a degree of synchronicity between a user action and its effect on the sound output with a short response, ultimately converging to real-time. The following figure represents a High-Level Model of a typical Virtual Music Instrument.

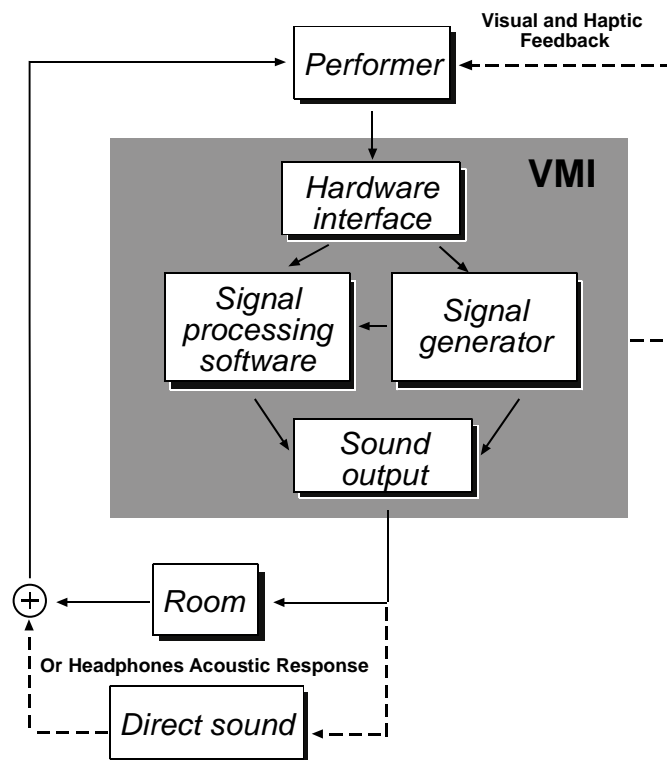


Figure 38. Virtual Music Instrument Interaction Model

In this model the hardware Interface can actuate on the Signal Generator typically to vary the Pitch and Dynamics as in a traditional instrument, but it can also actuate on the Signal processing software and, therefore, transform all kinds of parameters.

The signal processing software emphasis is usually on the spectral modeling of the signal allowing the user to change the timbre of the generated sound and also its pitch and dynamics. It is also possible to actuate on other parameters of sound like spatialization and duration at the processing stage. Many systems work only with one instance of the hardware interface actuating either in the signal generator (in many cases impersonating traditional instruments) or in the signal processing stage, in which case the signal generator is replaced by a sampler (feeding sound files with real world recordings to the Signal Processing Stage) or an input line from microphone capturing an acoustic sound.

In this diagram the sensorial feedback system related with the physical structure of the interface is also not represented in detail, but it is of great importance and most of the times will determine the success of the instrument. Yet, due to the experimental nature of this kind of devices the hardware interface is totally different from case to case. In a VMI the direct sound could be considered the output of the signal generator, which many times is not outputted from

the system. In most situations visual feedback from the system software is presented on the computer screen

### 3.2.2 Nomadic Music Instrument Model

In an interactive system designed to produce music, the sound synthesis engine and the user interface layer are fully integrated, but usually designed in parallel and in a modular way. Decoupling the interface layer from the synthesis engine, not only allows the use of best suited technologies and programming languages for each purpose, but also enhances overall system flexibility. In such a system architecture a remote user interface and a processing engine that resides in a different host can be taken to the most extreme situation in which a user can access the synthesizer from any place in the world using internet technology. This paradigm has promising applications in collaborative music creation systems for geographically displaced communities of users.

This paradigm complies to the notion of Sun's "Nomadic Computing" (Gadol, S. and Clary, M., 1994), introduced in the early 90's, i.e., where a network user moves and his familiar work environment must follow.

It should be noticed that the communication between the client and the server could have different content for each direction. In the Client-Server direction a real-time continuous connection is not as critical since often the communication is based on logic discrete commands triggered by the client interface. However, in the Server-Client direction an audio stream based on the sonic output of the synthesis engine is required to be transmitted as close to real-time as possible, therefore streaming and buffering techniques must be used (which increases latency). Another point that should be noted is that network delay introduced in both ways will be highly unpredictable and asymmetrical, since different content might be transmitted in each direction.

The following figure presents a model that defines this system through a perspective of an individual user.



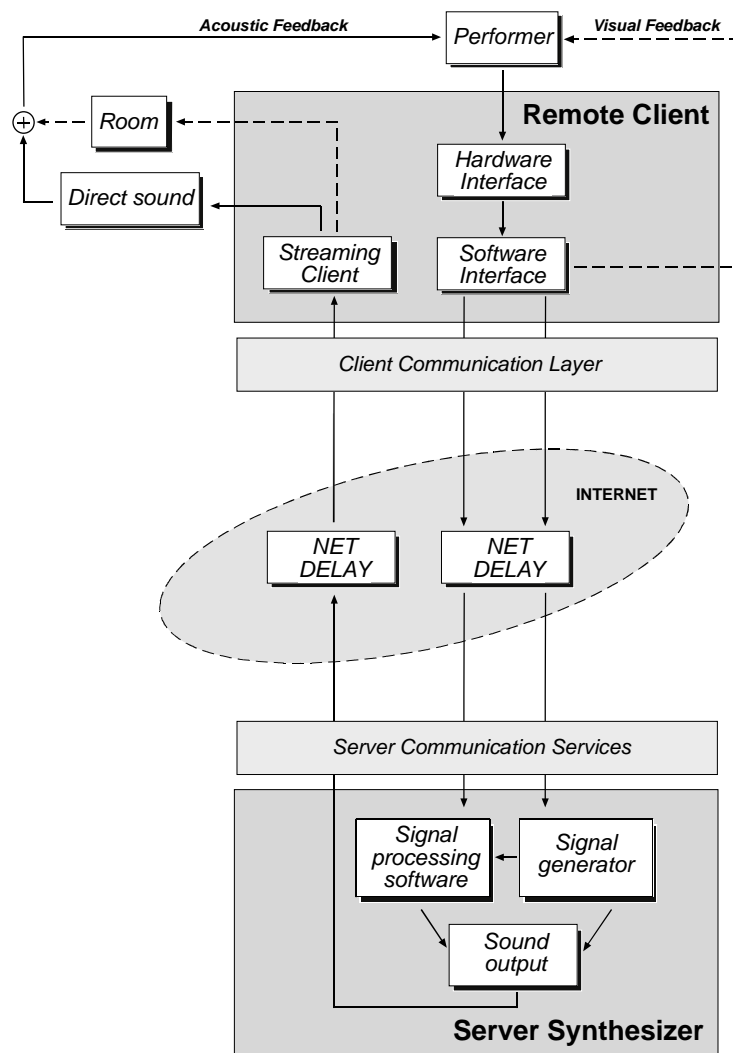


Figure 39. Nomadic Virtual Music Instrument Model

This model does not take into account the multi-user requirements, which would imply that the Signal Processing and Signal Generation stage would provide multi-user processing and that streamed output of the server synthesizer would convey the collective sonic performance of all the currently connected users. In this case local feed-back would have to be taken into account, as well as local synthesis and replication of interaction settings in order to obtain more accurate performances.

### 3.3 Multimodality and Networked Music

Music practices collaboratively performed by communities of users, especially on the Internet, have much to gain from a broader range of interactive mechanisms and devices accessible to the

common user. Mapping the human senses to technologically advanced devices is an ever-present activity in modern society. Because the sensory modalities are highly learned and natural, we try to find ways for machines to comply with this communication form. Technologies for image processing, voice interaction and tactile and manual interface control are burgeoning rapidly, but the dimensions of taste and smell are yet to be broadly addressed.

With computers Multimodality had an unprecedented deployment through experimental projects, even if the standard interaction paradigm between Human beings and computers is persistently based on a keyboard for discrete data input, a mouse for continuous and non-linear interaction and screen sound speakers for visual and auditory feed-back. Specifically in the case of Music Technology tactile interfaces have a major role in the design of musical instruments. Early examples of non traditional touch sensing applied to the control of electronic music devices are abundant. Much has been done over the last decades in extending the boundaries of traditional music instruments with electronic and digital technology, by researchers such as Tod Machover and Joe Paradiso at the Hyper-Instruments Group at the MIT Media Lab<sup>44</sup>. However, such developments date back to the early 20's with electronic music Instruments designed by Lev Termin, such as the well known "Theremin", the "Rhythmicon", the "Terpistone" or the "ThereminCello"<sup>45</sup>. More advanced interaction devices came up in the 50's, such as a large scale Touch-Bench designed by Oskar Fischinger in 1955 (Fischinger, O., 1955).

---

<sup>44</sup> More Information about the Hyper-Instruments Group is available from the MIT Media Lab website: <http://www.media.mit.edu/hyperins/>

<sup>45</sup> The earliest experiments with electronic Music Instruments can be traced back to 1897 with the registration of the "Telharmonium" Patent by the American Engineer Thaddeus Cahill's (1867-1934), However it was with the Russian cellist and electronic engineer Lev Sergeivitch Termen (1896- 1993) that new interaction methods, inherent to the emergent possibilities of electronic Technology (the vacuum-tubes era), came into Instrumental Music Practice. More information is available from: [http://www.obsolete.com/120\\_years/](http://www.obsolete.com/120_years/)

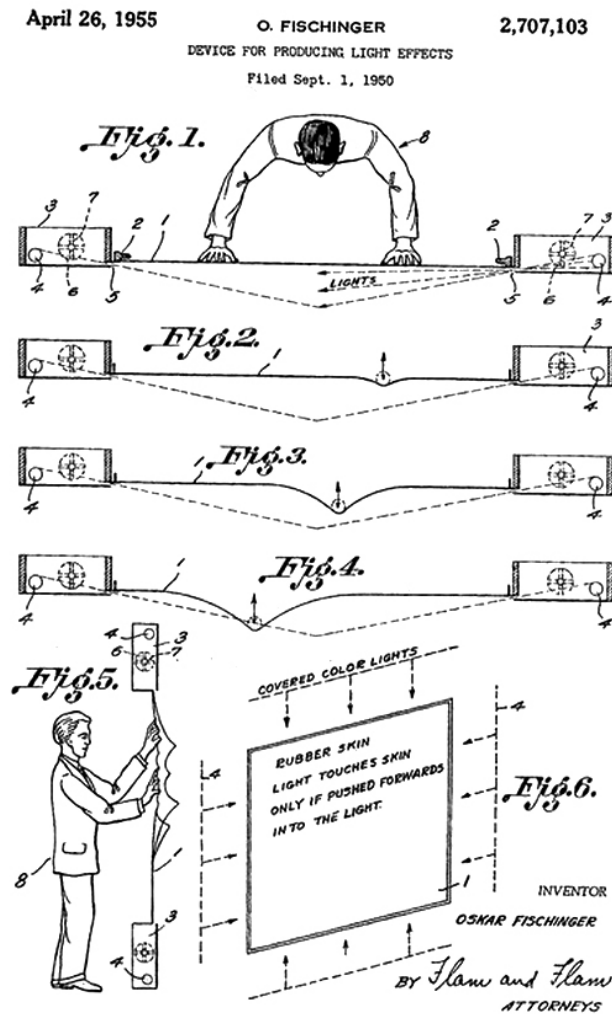


Figure 40. Oskar Fischinger's device for producing light effects.

Similar large scale Touch-Screens are presently very common as controllers for musical instruments based on computer technology, with the major difference that such devices are now affordable to be built by a much wider spectrum of users. This is due to the financial accessibility of technology and generalized access to information.

An example of similar technology in 2005 is a simple technique that enables robust multi-touch sensing at a minimum of engineering effort and expense.

It relies on frustrated total internal reflection (**FTIR**). It acquires true touch information at high spatial and temporal resolutions, and is scalable to very large installations<sup>46</sup>.

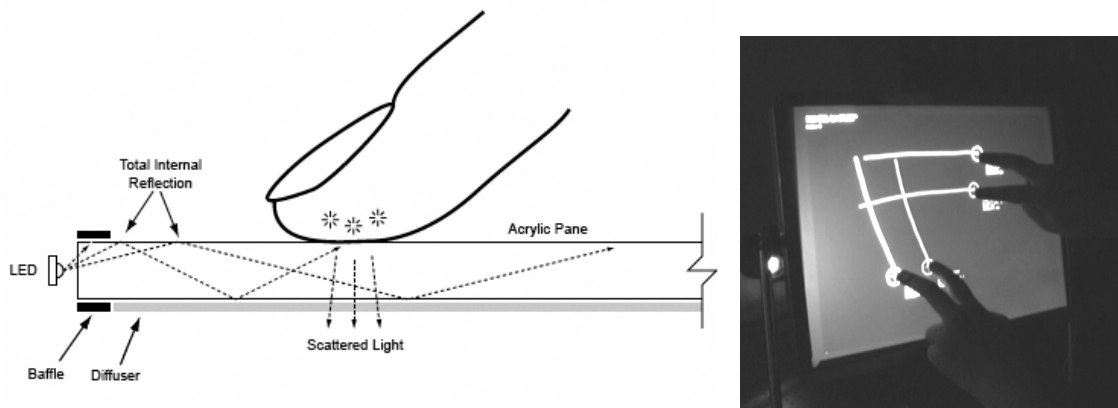


Figure 41. Multi-Touch Sensing through Frustrated Total Internal Reflection

In fact at the present moment some input devices can be considered as broadly accessible to the majority of computer user's community, such as microphones (for audio input), web cams (for video input) and mobile touch screens available from Personal Digital Assistants (PDAs).

### 3.3.1 Personal Digital Assistants (PDAs) as Music Controllers

PDAs or Pocket PCs provide unique features that are of great relevance for virtual music instruments.

- Wireless Mobility
- Reasonably powerful computation capacity
- Touch-Screen interaction (overlaps visual feed-back with continuous linear control)

These attributes are newly gained into the general interaction facets that can be expected to be accessible to a common computer user, and therefore are of great importance in an ubiquitous musical communication paradigm.

---

<sup>46</sup> More information about Multi-Touch Sensing through Frustrated Total Internal Reflection is available from: <http://mrl.nyu.edu/~jhan/ftirtouch/> (Accessed 3 November 2005)

In 2001, Gunter Geiger started a project that brought music performance potential to PocketPCs (Geiger, G., 2003). The project consisted of porting the computer music systems Pure Data (PD) (Puckette, M., 1996a, Puckette, M., 1996b) to a PDA platform. PD shows a great extensibility and was originally designed to run on SGI Irix and Windows NT. It was ported to Linux in 1997 and runs on all major desktop operating systems. This was only possible because of its openness and availability as free software.



Figure 42. A Compaq IPAQ Running a PD Patch, Gunter Geiger performing with a PDA

A different approach was used in the Public Sound Objects Project, discussed later in chapter 5. In this case a PDA was used mainly as a client interface developed in Java (an applet running on a web browser at the Pocket PC), since the computation required for sound synthesis was too demanding to be processed by a PDA.



Figure 43. The Public Sound Objects Client Interface running on a PDA

The client applet at the PDA communicated through the internet with a Server, sending control data and receiving an audio stream resulting from the user's action on the Touch-Screen.

### 3.3.2 The ReacTable\*

**ReacTable\*** is an Electronic Musical Instrument developed at the Interactive Systems Group at the MTG. This project led by Sergi Jordà was presented at ICMC 2005 as follows:

“The reacTable\* is a state-of-the-art music instrument, which seeks to be collaborative (local and remote), intuitive (zero manual, zero instructions), sonically challenging and interesting, learnable and masterable, and suitable for complete novices (in installations) and for advanced electronic musicians (in concerts). The reacTable\* uses no mouse, no keyboard, no cables, no wearables. The technology it involves is, in other words, transparent to the user; it also allows a flexible number of performers that can enter or leave the instrument-installation without previous announcements.” (Jordà, S. and others, 2005)

ReacTable\* has been a central project at the MTG, since it conveys in one musical applications several development efforts conducted by researchers at the interactive systems group.

The instrument is based on a translucent round table, with a video camera positioned underneath, continuously scanning the table surface and tracking the nature, position and orientation of the objects that are distributed on it. The objects are passive and of different shapes, without any sensors or actuators. Users interact by moving them, changing their position, their orientation or their faces (in the case of volumetric objects), controlling with these actions the topological structure and the parameters of a sound synthesizer. Also from beneath the table, a projector draws dynamic animations on its surface, providing a visual feedback of the state of the synthesizer.

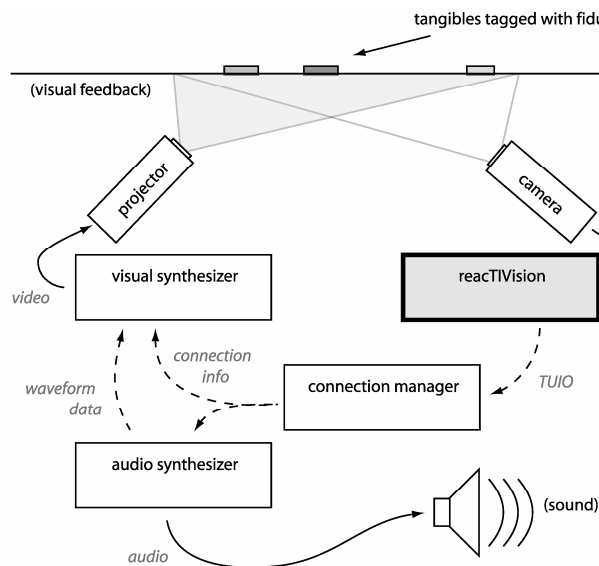


Figure 44. The ReacTable\* architecture (illustration by Ross Bencina)

In order to produce this system custom made software had to be developed, such as, dynamic patching capabilities module for Pure-Data required for the sound engine (Kaltenbrunner, M., Geiger, G. and Jordà, S., 2004), and a Fiducial Tracking system for the computer vision engine (Bencina, R., Kaltenbrunner, M. and Jordà, S., 2005). From this complex and time-consuming research effort results a tangible musical instrument that is simultaneously inexpensive and reasonably straightforward to build, and therefore accessible to a very wide community of potential users.

A musical composition for this instrument was commissioned to the composer Chris Brown, and the resulting piece “**TeleSon**” was premiered simultaneously (performed in remote

collaboration over the Internet) at the International Computer Music Conference 2005 in Barcelona and the Ars Electronica Festival in Linz.

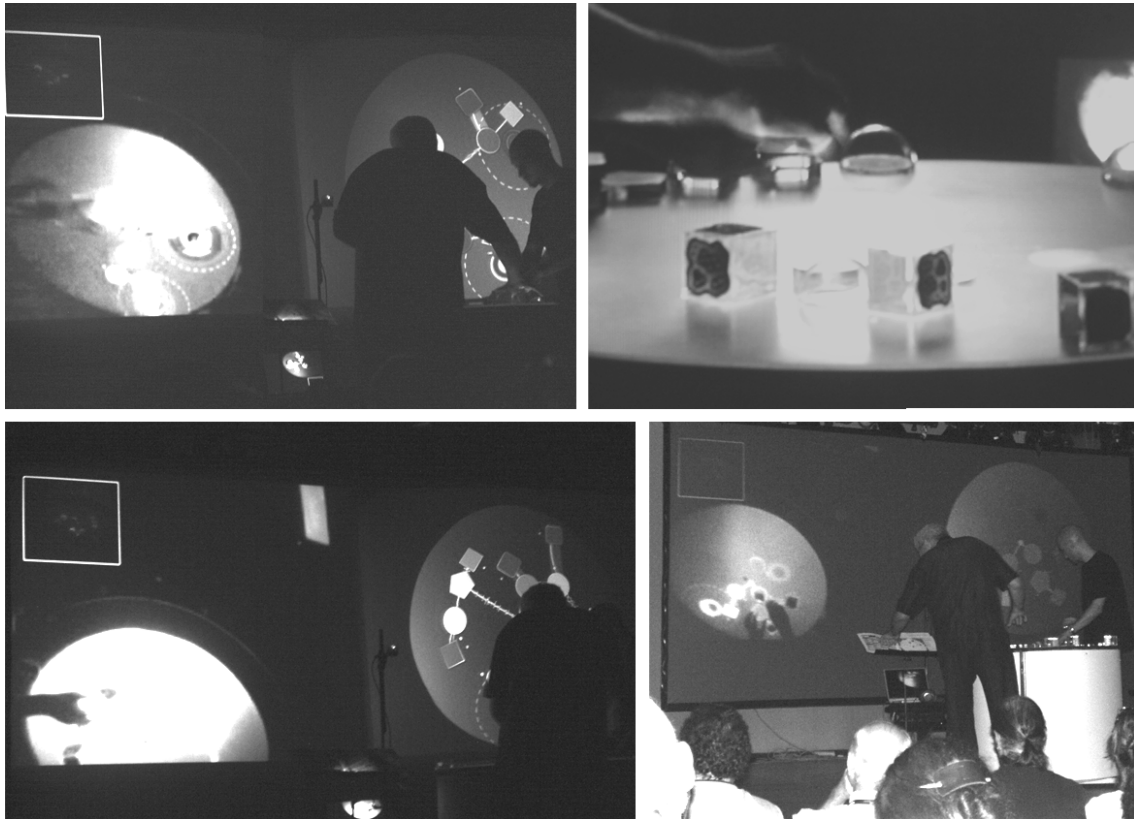


Figure 45. “TeleSon” Performance September 04, 2005: Chris Brown and Gunter Geiger at ICMC 2005 in Barcelona, Spain (on stage at SGAE auditorium); Martin Kaltenbrunner and Marcos Alonso at Ars Electronica Festival in Linz, Austria (on screen).

“TeleSon” clearly outlines the collaborative model developed specifically for ReacTable\*. It provides a very strong co-located collaboration paradigm that can also be applied in geographically displaced scenario, preserving all its interaction and interconnection



features.

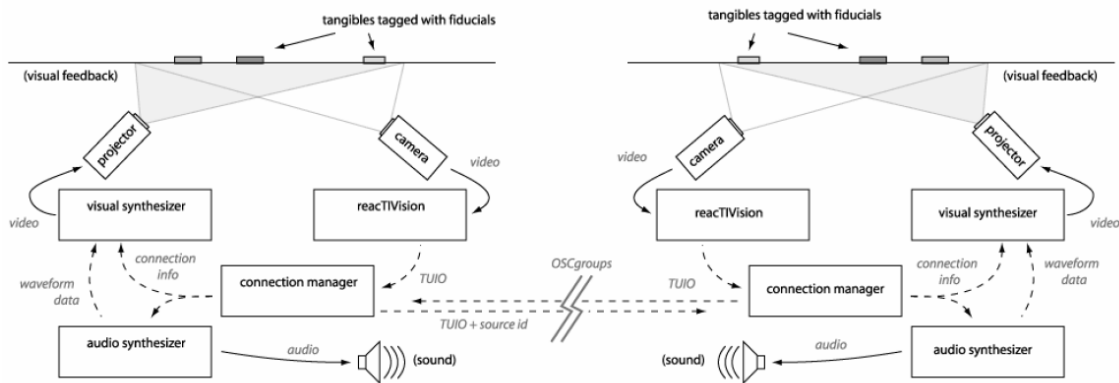


Figure 46. Schematics for two Networked ReactTables\* (illustration by Ross Bencina)

Once again in order to implement a smoothly working prototype of Networked ReactTables\* new technology had to be researched and developed. In 2005 during his collaboration with the MTG, Ross Bencina wrote an extension to the communication protocol Open Sound Control entitled **OSCGroups**.

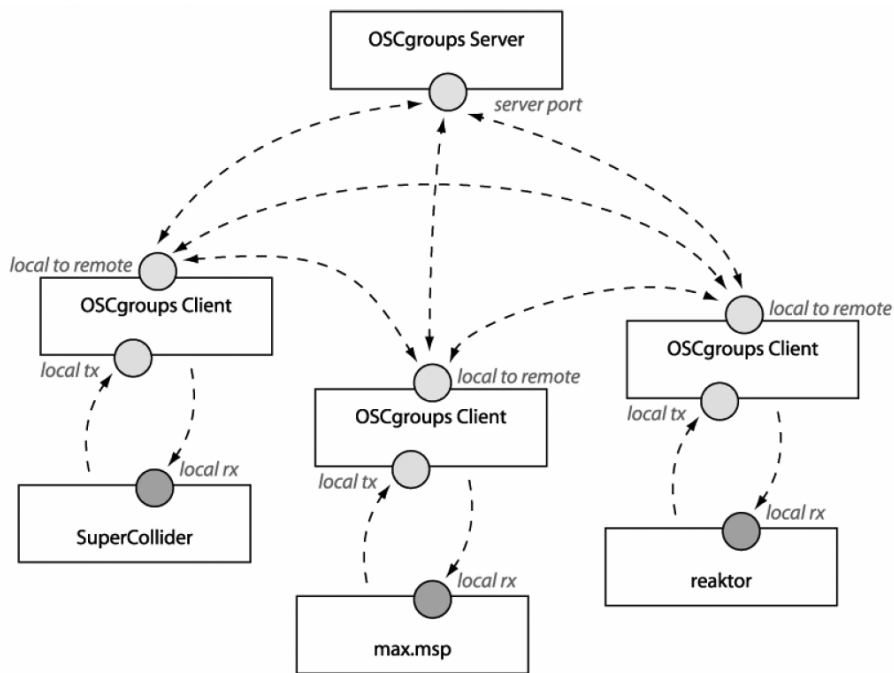


Figure 47. OSCGroups Communication Model by Ross Bencina

Bencina addressed the main issues raised by data transmitting through OSC over the internet: Packet loss and reordering; Router/firewall Network Address Translation (NAT); Peer discovery (naming services).

OSCGroups handles NAT issues with *NAT Traversal* (hole punching)<sup>47</sup>, manages group membership and allows client to use peer-to-peer multicast to other group participants.

The Group Membership Model provides separate password Users and Groups, similarly to IRC. If a user doesn't stay connected someone else can use his name after a timeout period, otherwise they need his password.

OSCGroups is unquestionably a major development for Networked Music and, in fact, Groups like the Hub are already using it for remote Musical rehearsal and performance.

## 3.4 Chapter Conclusions

The idea of having an interface and a sound synthesizer accessible over a computer network, without geographical constraints, is a paradigm presented in this chapter as a **Ubiquitous Virtual Music Instrument**. There are different approaches and models in order to implement a collaborative music system, as presented in Chapter 2<sup>48</sup>, of this dissertation, and it is a relatively new and unexplored research topic, as presented on section 2.2.3 of the same chapter. However, when implementing such systems one can refer to Computer Science concepts and methods and articulate it with recent research in the context of Virtual Music Instruments. Network protocols, services and models are presently text-book material for any Telecommunications Engineering graduate degree and new Musical instruments designed with digital technology are a central research topic at a Doctoral level in Music Technology (Jordà, S., 2005a). In such a framework it is possible to define high level models for general cases of networked music practice, which can combine different topologies based on centralized, distributed and peer-to-peer models, in order to attend to Networked Music System' requirements.

---

<sup>47</sup> Accessible from: Bryan Ford, Pyda Srisuresh and Dan Kegel, "Peer-to-Peer Communication Across Network Address Translators," <http://www.brynosaurus.com/pub/net/p2pnat/>, p5.

<sup>48</sup> Refer to section 2.4 - Networked Music Systems Overview

More specifically it is presented a general model designed to support “Interface Decoupled Applications for Geographically Displaced Collaboration in Music” (Barbosa, A., Kaltenbrunner, M. and Geiger, G., 2003), the **Nomadic Music Instrument Model**, which served as a reference framework for Web Services and Acoustic Applications, as well as, the Public Sound Objects System presented respectively on Chapter 4 and 5 of this dissertation.

Furthermore, a brief introduction to the issue of until what extent multimodality can be incorporated in Networked Music practice in the era of technology accessibility. The focus is on the Personal Digital Assistant (PDA) as client interface for Musical Applications and the ReactTable Project as multi-user instrument developed with free software and publicly accessible technology and materials.

# Chapter 4

## Internet Acoustic Communication Facets

“I find Internet time delay rather interesting and I think of it as a kind of unique acoustic of this media (...) rather than to play exiting music on this new time basis, what is interesting to me is trying to find a musical language that works on this time axis (...) if it takes half a second of delay for a sound to go from Paris to New York and another half a second to come back, then we can create a music that is adapted to this acoustic (...).”  
(From Atau Tanaka’s Video Interview at Golo Föllmer’s essay “Soft Music” (Föllmer, G., 2001))

The advent of internet computing and the possibility of acoustic communication over IP brought the opportunity of geographically displaced musical performance to a worldwide community. However, it is well known that Network Latency (or net-delay) has a highly disrupting effect in this practice, especially in traditional music performance, driven by rhythm and melody, requiring very tight synchronicity to achieve a desirable real-time mutual awareness amongst participants.

The thought expressed by Tanaka, that Internet time delay is the unique acoustics of Internet and that composers should create music embracing this fact, is an inspiring view of this topic. It somehow brings up the recurrent notion that adapting music to the media where it is performed leads to stylistic novelty, such as it was stated the example on section 2.3.1 from chapter 2 about Venetian polychoral music style originated from the peculiar acoustic space of the Basilica San Marco di Venezia in Italy.

In this chapter is presented research related with the **use of internet from a musical perception perspective** and examples of **web services and applications** developed based on results from this research and the network music practice topologies presented in chapter 3.

## 4.1 The Perception of Internet Acoustics

The SoundWire group at the Center for Research in Music and Acoustics (CCRMA) at Stanford University, lead by Chris Chafe, published several research articles over the last few years referring to the implications of network conditions on acoustic communication (Chafe, C. and others, 2000) (Chafe, C. and Leistikow, R., 2001).

Specifically in the article “*Physical Model Synthesis with Application to Internet Acoustics*” (Chafe, C., Wilson, S. and Walling, D., 2002), Chafe describes how distributed physical models of musical instruments have been used to acoustically “ping” Internet connections between two network hosts, departing from the observation that sound waves propagated through Internet acoustics behave just as in air, water or along a stretched string.

The idea of “listening to the sound of a network” is a stimulating view of how Network latency can be regarded as a major characterizing feature of Internet’s Acoustics.

In collaboration with the artist Greg Niemeyer, this same idea led to the experimental music installation at San Francisco Museum Of Modern Art entitled Ping (Chafe, C. and Niemeyer, G., 2001), in which synthesis by physical models is used for internet data sonification.

Internet has different facets that can have an effect on any collaborative process. Real-time synchronicity is unquestionably central in Musical practice, and latency is a major drawback for real-time music collaboration in general.

This problem is present in many other contexts besides communication over long distance computer networks, such as in computer sound cards or in audio amplification systems in large auditoriums where the sound coming out from the rear speakers needs to be electronically delayed to match the phase of incoming stage sound delayed during its propagation through the atmosphere.

It is commonly presented, as an illustrative example of the disrupting effect of acoustic delay caused by the propagation of sound in the atmosphere, a scenario where two musicians try to play together placed on opposite sides of a foot-ball stadium (about 120 meters apart). The sound will take about 35 ms (considering Sea level, 15 °C – 340 m/s Speed of Sound) to go from one musician to the other, and the round-trip auditory Feed-Back, for each musician to be aware of the others reaction, would be twice that value (70 ms). These values are very high and it would be very hard to achieve a smooth and synchronized performance.

### 4.1.1 Latency Tolerance in Music Performance

For the Human ear to perceive two simultaneous sounds, they should not be displaced in time over 20ms (Hirsh, I., 1959), which means that for mutual awareness to be supported in a bilateral performance this threshold would be around 40ms (the time period that it would take one performer to perceive the other performers reaction to his action).

It should be noted that the perception of two different sounds performed simultaneously is strongly dependent on sound characteristics (timbre, pitch or loudness), music style and other feedback types, such as visual or physical stimuli. Nevertheless, a 20 ms threshold is reasonably high enough to characterize a worst-case scenario.

In fact, a number of experiments were carried out with the purpose of determining the maximum amount of communication latency which can be tolerated between musicians in order to keep up with a synchronous performance.

Significant results from research carried out in 2002 at Stanford University by Natham Shuett (Schuett, N., 2002) established experimentally an **Ensemble Performance Threshold (EPT)** for impulsive rhythmic music lying between 20-30 ms, which is consistent with the outcome from research carried out by Nelson Lago in 2004 (Lago, N. and Kon, F., 2004) at São Paulo University.

In the context of audio transmission over Computer Networks, considering advances in Broadband and Compression performance one could be led to believe that Network Latency is a Technological condition that can be overcome in the near future, and therefore it would be somewhat useless to study forms to diminish its disturbing effect in traditional musical performance.

Even if we do not consider the extreme latency introduced in satellite communication, or that emergent Mobile Technology has much slower data transfer rates, it can be demonstrated by the laws of Physics that at a globe level there are limits, which will always introduce values of latency higher than the minimum acceptable threshold for real-time musical collaboration.

If we consider the smallest possible peer-to-peer connection between two opposite points on the planet, lets say, Santiago do Chile and Moscow, we have an approximate distance of 14.141 Km. Even with ideal data transfer at the speed of light (299.792,458 Km/s) and unlimited bandwidth,

bi-directional latency would be approximately 94,3 ms, which is much higher than a minimum tolerable threshold.



Figure 48. Ideal Communication Scenario between two globally displaced Cities.

Furthermore, Latency has a highly variable and unpredictable nature creating time base errors, de-sequencing and even partial loss of the content, resulting in a severe condition for performance control. Nevertheless, a major effort is being made by the scientific community, in order to diminish this condition, by increasing bandwidth, compressing data and by using content based transmission techniques.

Therefore, if one considers Large Area Networks or even Wide Area Networks in geographically constrained territories (a country or even a continent) it can be expected that in the near future network latency is likely to be reduced to values which will not represent an impediment for real-time acoustic communication over the internet.

### 4.1.2 Latency Adaptive Tempo and Dynamics

Some research concerning the effects of time delay on ensemble accuracy goes beyond establishing an EPT for a general scenario of rhythmic synchronization.

The work published in 2004 by Chris Chafe and Michael Gurevish (Chafe, C. and others, 2004), resulting from an experiment conducted at CCRMA, shows that by increasing the

communication delay between pairs of subjects trying to synchronize a clapping steady rhythm, the subjects tend to slow down the rhythm tempo.

Similarly, an experiment carried out by the author of this dissertation in June 2004 at the Sound and Image Department of the Portuguese Catholic University (UCP) aimed, amongst other goals, to study the relationship between Tempo and Latency. In the experiment, simulated network latency conditions were applied to the performance of four different musicians playing jazz standard tunes with four different instruments (Bass, Percussion, Piano and Guitar).



Figure 49. Experiment on latency tolerance in a simulated studio environment

The first part of this experiment consisted in determining the maximum **individual latency tolerance** applied to the auditory feed-back from the musicians own instrument.

For this purpose a studio system was set up, so that musicians would listen to the feed-back from their own instruments through headphones with variable delays. Performances were synchronized with a metronome over several takes with different tempos (Beats Per Minute – BPMs). For each take the feed-back delay was increased until the musician wasn't able to keep up a synchronous performance.

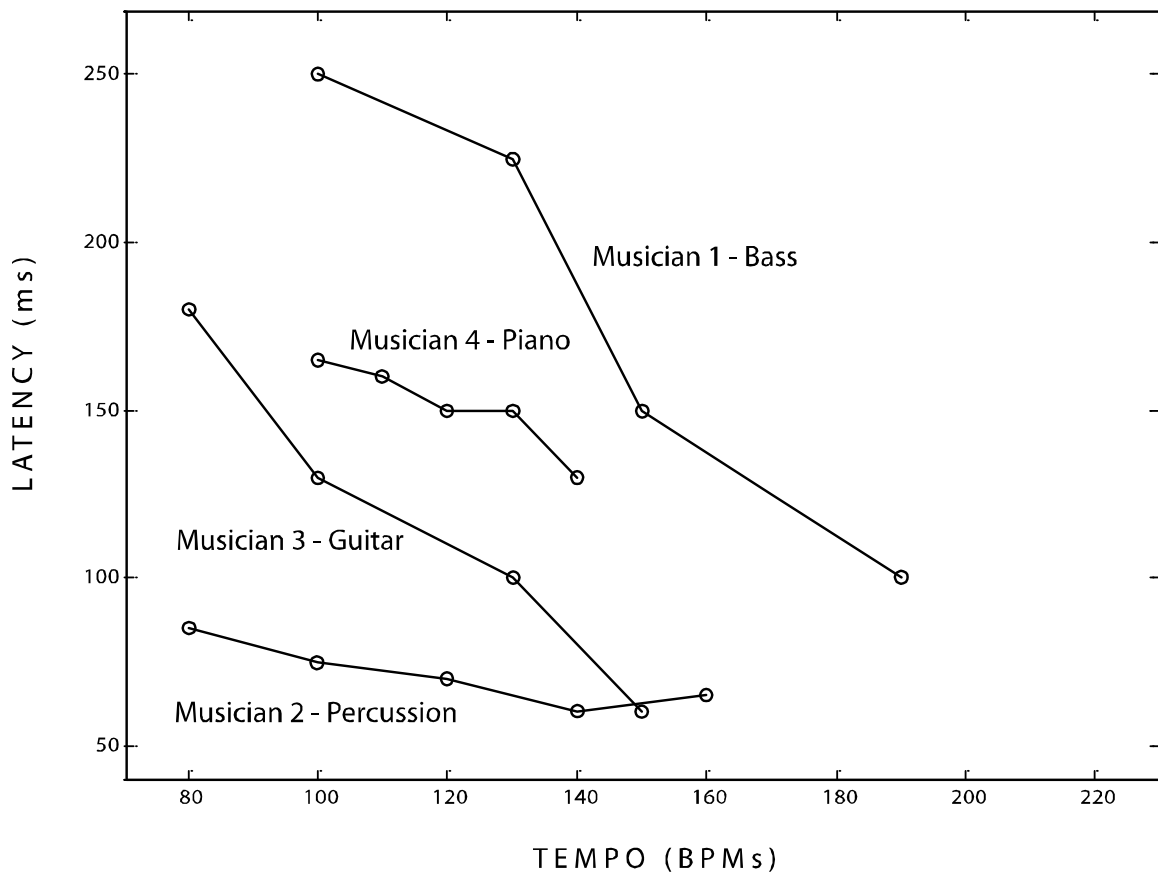
The following graphic and table show the results from this preliminary experiment.

		<b>Musician</b>			
		<b>Bass</b>	<b>Percussion</b>	<b>Guitar</b>	<b>Piano</b>
80	(BPMs)	-	85 ms	180 ms	-



<b>100</b>	250 ms	75 ms	130 ms	165 ms
<b>110</b>	-	-	-	160 ms
<b>120</b>	-	70 ms	-	150 ms
<b>130</b>	225 ms	-	100 ms	150 ms
<b>140</b>	-	60 ms	-	130 ms
<b>150</b>	150 ms	-	60 ms	-
<b>160</b>	-	65 ms	-	-
<b>170</b>	125 ms	-	-	-
<b>190</b>	100 ms	-	-	-

Table 1. Maximum delay tolerance for each musician playing at different tempos



Graphic 1. Self-Test for latency tolerance in individual performance

It is clear that regardless of instrumental skills or music instrument, all musicians were able to tolerate more feed-back delay for slower Tempos.

The only exception to the main decreasing tendency of these curves occurs for the percussionist when raising to 160 BPMs, which is related to a synchronous overlap over the music rhythmic structure, together with the fact that with percussion instruments it is very hard to totally isolate the performer from direct instrument sound. Therefore, it is reasonable to assume that **there is an inverse relationship between Musical Tempo and Latency-Tolerance.**

For further evaluation of this assumption a user test was conducted in the context of this experiment. An on-line survey was submitted to 32 subjects with a dominant profile of music student's from the School of Arts of the Portuguese Catholic University (53% with academic music training; 28% can play a musical instrument; 19% without any musical training).

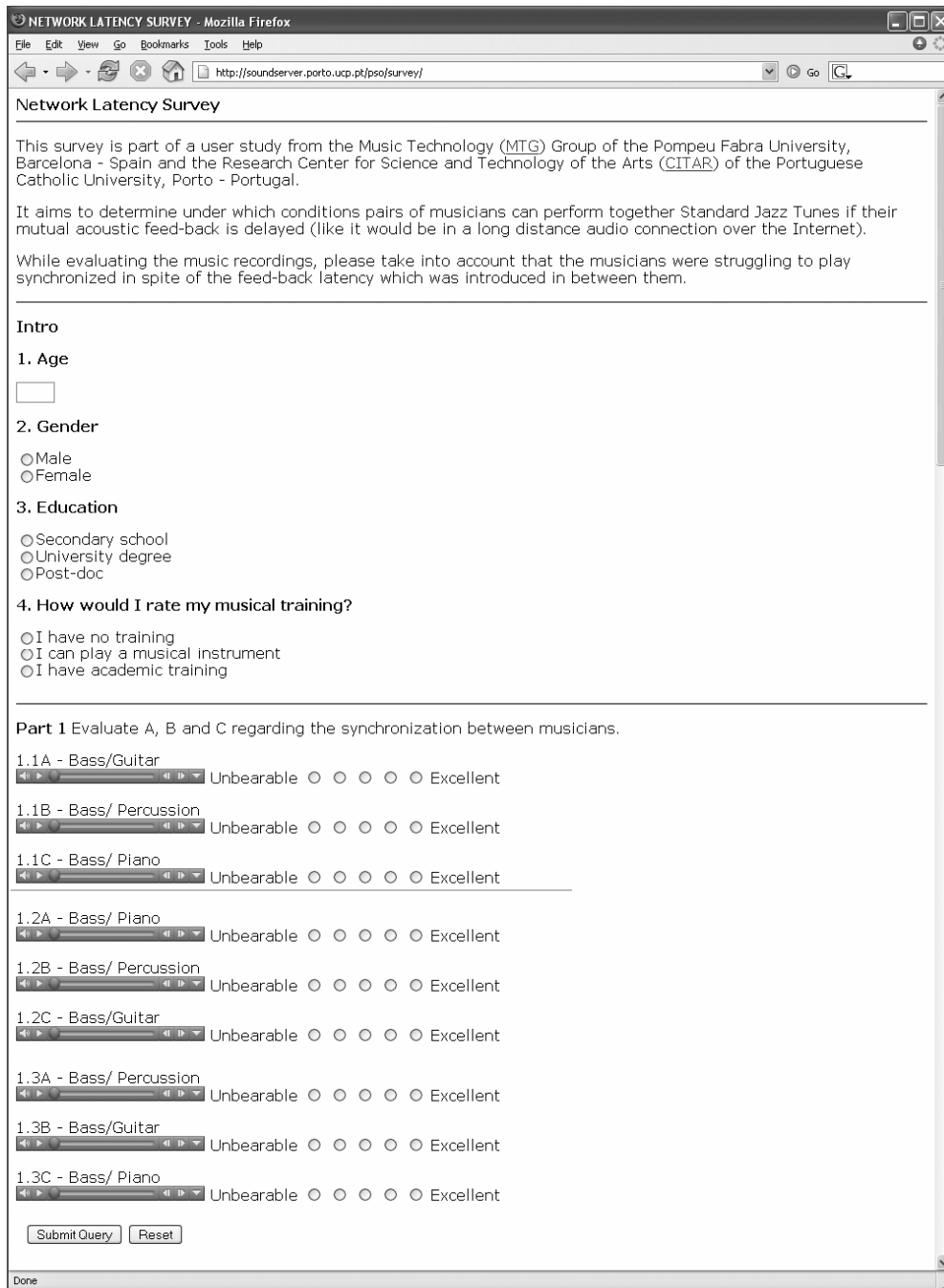


Figure 50. Online survey to evaluate the relationship between Musical Tempo and Communication Latency

The survey consisted of asking for a classification of performance accuracy of the well known Jazz Song “Sunny” written by Bobby Hebb, by different pairs of instruments (Bass/Percussion; Bass/Guitar; Bass/Piano) and for different combinations of Tempo (BPMs) at a fixed communication latency between musicians of 30 ms (35 ms in case of the Bass/Piano duet).

**C INSTRUMENTS - "SUNNY" 145.**

**BRIGHT** ♩ = 138

Figure 51. Music notation for the song “Sunny” from the Jazz Real-Book

The song “Sunny” was preferred since it was one of the songs performed in the Stanford-McGill experiment (June 13, 2002: McGill-Stanford trial)<sup>49</sup>, in which it was clear from empirical observation of the video documents<sup>50</sup> that, in some moments, musicians could not keep good performance synchronization. This way recorded material in Porto’s sessions could be compared in future research with the Stanford-McGill experiment.

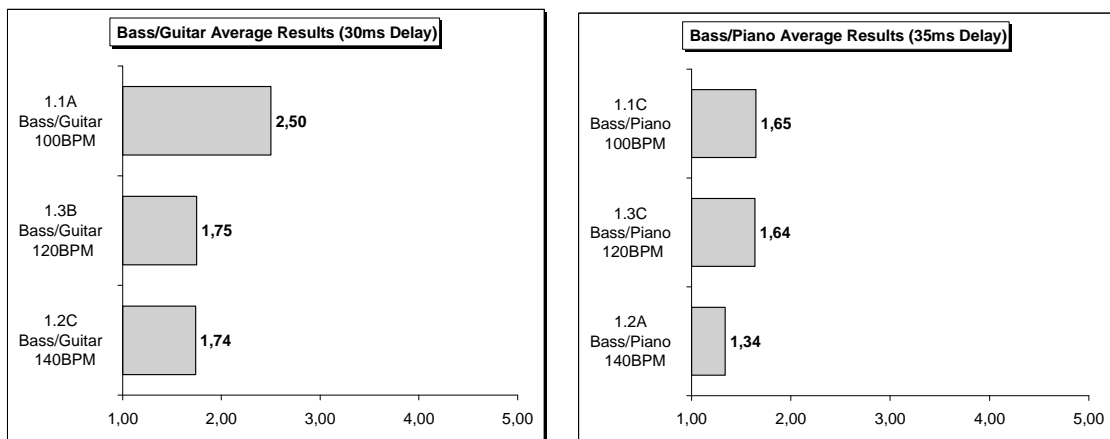
Results presented in the following tables and graphics show that most subjects considered that performances with approximately the same latency are generally better for lower Musical Tempo (100 BPM), regardless of the instruments and performers, which confirms the initial hypothesis.

<sup>49</sup> Discussed in Chapter 2, Section 2.4.3.2

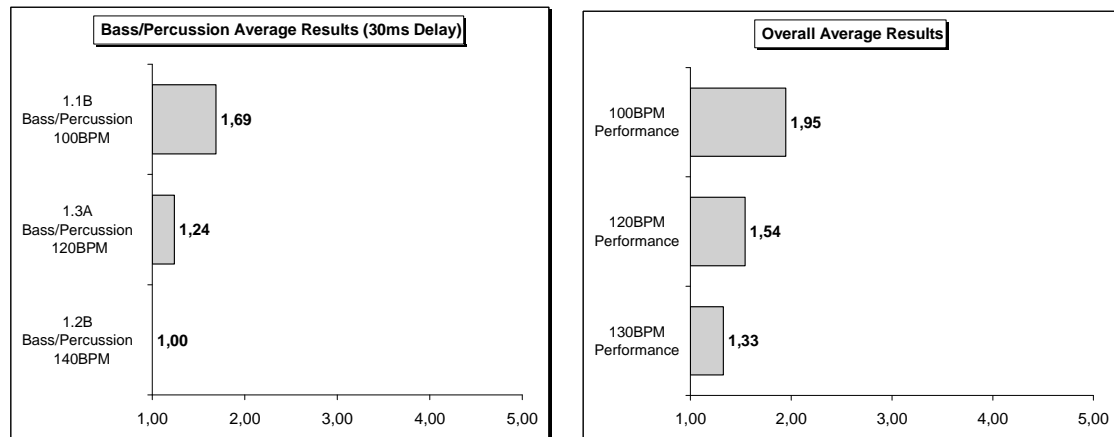
<sup>50</sup> Available from: <http://www.cim.mcgill.ca/sre/projects/rtnm/>

Date	IP Address	Age	Gender	Education	Musical Training	1.1A	1.1B	1.1C	1.2A	1.2B	1.2C	1.3A	1.3B	1.3C
27-04-2005 11:40	172.20.80.60	20	M	University degree	Academic training	3	0	2	1	0	1	0	2	1
20-04-2005 14:58	172.20.80.60	25	F	University degree	Academic training	3	0	2	0	1	3	0	2	0
19-04-2005 12:07	172.20.80.60	38	F	University degree	Academic training	2	1	3	0	0	2	0	1	2
29-06-2005 16:19	172.20.80.60	31	M	University degree	Academic training	0	0	3	0	0	1	2	0	1
18-04-2005 16:24	172.20.80.60	23	M	University degree	Academic training	4	1	3	0	0	2	2	2	3
18-04-2005 16:20	172.20.80.60	26	M	University degree	Academic training	3	1	2	1	0	3	1	0	3
02-05-2005 16:25	172.20.80.60	25	F	University degree	Academic training	3	1	1	1	2	1	2	2	0
02-05-2005 16:58	172.20.80.60	25	M	University degree	Academic training	2	1	0	1	0	1	1	0	1
03-05-2005 15:00	172.20.80.60	23	F	University degree	Academic training	2	3	3	2	1	2	0	2	1
03-05-2005 17:26	172.20.80.60	32	M	University degree	Academic training	2	1	2	3	1	2	2	2	1
06-05-2005 12:37	172.20.80.60	28	M	University degree	Academic training	1	0	0	1	0	2	2	0	1
07-05-2005 11:57	84.143.179.74	45	M	Post-doc	Academic training	1	2	0	1	1	2	1	0	1
11-05-2005 19:17	193.145.55.204	28	M	University degree	Academic training	3	1	0	2	1	0	1	1	0
11-05-2005 20:32	193.145.55.204	29	M	University degree	Academic training	2	1	3	1	2	2	3	2	2
12-05-2005 14:59	172.20.80.60	24	F	University degree	Academic training	2	0	2	1	1	0	1	2	0
19-05-2005 10:45	172.20.80.60	28	M	University degree	Academic training	2	0	1	0	1	3	1	1	0
24-05-2005 11:18	172.20.80.60	24	F	University degree	Academic training	3	2	2	1	0	2	1	2	2
Average:						2,24	0,88	1,71	0,94	0,65	1,71	1,18	1,24	1,12
28-04-2005 10:09	172.20.80.60	23	M	University degree	Can play musical instrument	3	1	3	0	2	2	3	2	2
28-04-2005 10:34	172.20.80.60	26	M	University degree	Can play musical instrument	3	2	3	1	2	3	2	1	2
28-04-2005 11:55	172.20.80.60	34	M	University degree	Can play musical instrument	3	1	3	0	1	2	1	2	2
03-05-2005 10:14	194.117.24.10	31	M	University degree	Can play musical instrument	3	0	1	3	1	2	2	2	1
03-05-2005 13:07	172.20.80.60	36	F	University degree	Can play musical instrument	3	3	1	1	0	2	1	1	3
07-05-2005 17:46	172.20.80.60	24	M	Secondary School	Can play musical instrument	3	2	2	3	1	1	0	3	1
11-05-2005 18:07	193.145.55.204	37	M	University degree	Can play musical instrument	2	1	2	1	1	1	1	0	1
11-05-2005 20:03	193.145.55.204	25	M	University degree	Can play musical instrument	1	2	1	1	1	1	1	2	1
16-05-2005 15:15	141.83.78.62	28	F	University degree	Can play musical instrument	4	1	2	2	1	2	3	3	1
Average:						2,78	1,44	2	1,33	1,11	1,78	1,56	1,78	1,56
23-04-2005 16:08	172.20.80.60	33	M	University degree	No Training	3	3	2	4	1	1	1	2	0
30-04-2005 17:54	192.35.246.5	36	M	University degree	No Training	1	2	0	0	0	2	1	1	3
02-05-2005 11:19	172.20.80.60	34	F	University degree	No Training	3	3	2	1	3	0	0	4	4
10-05-2005 11:23	193.145.56.194	26	F	University degree	No Training	3	3	1	2	1	4	2	2	2
Average:						2,5	2,75	1,25	1,75	1,25	1,75	1	2,25	2,25
Final Average:						2,5	1,69	1,65	1,34	1	1,74	1,24	1,75	1,64

Table 2. Evaluation results for Musical Tempo/Communication Latency relationship grouped by musical training levels



Graphic 2. Evaluation results for Musical Tempo/Communication Latency relationship in the case of Bass/Guitar and Bass/Piano duets



Graphic 3. Evaluation results for Musical Tempo/Communication Latency relationship in the case of Bass/Percussion duet and a final overall average result

Direct dependency between musical Tempo and tolerance to the disrupting effect of latency in this specific case of music collaboration (standard Jazz performance), can be regarded as a more general concept of **Latency Adaptive Tempo (LAT)**.

The basic application principle of LAT consists of a simple software function, for network acoustic communication systems, that dynamically adapts the Musical Tempo (typically a referenced by a metronome sound) to the maximum value tolerated by the least “latency-tolerant” musician of an ensemble. This dynamic adaptation is based on real-time latency measurement between peers.

Input variables of this function are musicians’ profiles and latency value at a given moment. The output of the LAT function will be the Tempo value (typically in BPMs) that is less disrupting for the group musical practice.

LAT allows musicians to rehearse music as fast (in terms of Musical Tempo) as their Network connectivity speed allows them to.

This concept was implemented into the Public Sound Objects (PSOs) system, discussed in chapter 5, with respective adjustments to the Musical Tempo concept and latency-tolerance requirements of this particular Networked Music instrument.

Furthermore it should be acknowledged that, from a different perspective, the idea of a network music instrument which dynamically adapts to internet network-latency was

implemented recently by Jörg Stelken in the peerSynth Software (Stelkens, J., 2003). PeerSynth is a per-to-peer sound synthesizer<sup>51</sup> which supports multiple users displaced over the internet, measuring the latency between each active connection and dynamically lowering the sound volume of each user's contribution on the incoming soundscape, proportionally to the amount of delay measured in his connection. Stelkens followed a real world metaphor where, in fact, the sound volume of a sound source decreases with the distance to the receiver, which also implies increasing acoustical communication latency. A similar approach was followed in the AALIVENET System (Spicer, M., 2004).

### 4.1.3 Individual Delayed Feed-Back

Another result that came out of the experiments with simulated acoustic communication latency at the School of Arts of the Portuguese Catholic University was a Feed-Back Topology that enhances individual latency-tolerance.

It was observed empirically that, with practice, musicians tend to improve their skill to play their musical instrument when their individual acoustic feed-back suffers delay. This idea is reinforced by the results presented in Graphic 1, in which we can observe different levels of tolerance to individual feed-back for musicians with different instrumental skills.

This also led to the belief that better latency tolerance is achieved if instead of having ensembles in which each musician receives direct acoustic feedback from their own instrument mixed with delayed feedback from the other performers, every musician listens to his acoustic feed-back delayed, together and in sync with the others. This concept is defined as **Individual Delayed Feedback (IDF)**.

The following figures illustrate the experimental studio setup used for recording of sessions between pairs of musicians. The same song was recorded with the same tempo and latency, but using normal Feedback Topology in one take and an Individual Delayed Feedback Topology in another take.

---

<sup>51</sup> Discussed in Chapter 2, Section 2.4.3.2

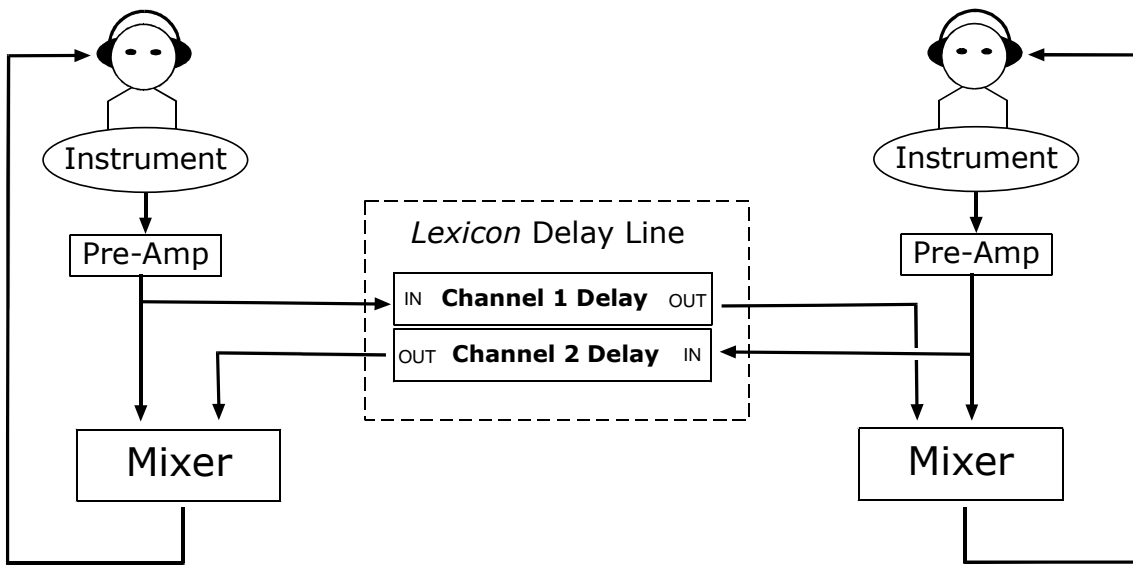


Figure 52. Normal Feed-Back Topology

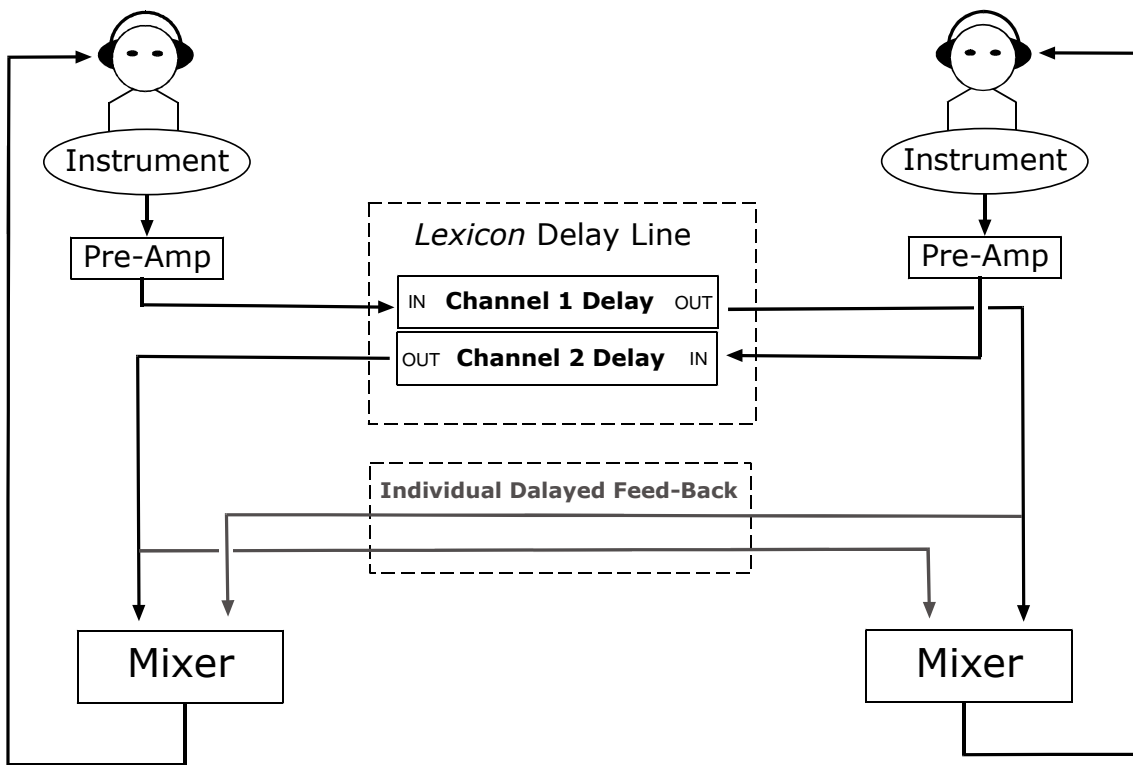


Figure 53. Individual Delayed Feed-Back Topology

The song recorded with these two Feed-Back Topologies was the well known jazz tune “Cantaloupe Island”, written by Herbie Hancock, at the tempo of 120 BPM for a communication Delay of 35 ms.





Figure 54. Music notation for the song “Cantaloupe Island” from the Jazz Real-Book

Four different pairs of instrumental performances were recorded: Bass/Guitar; Bass/Percussion; Bass/Piano; Piano/Percussion. These Recordings were used in the following On-Line User Survey<sup>52</sup>:

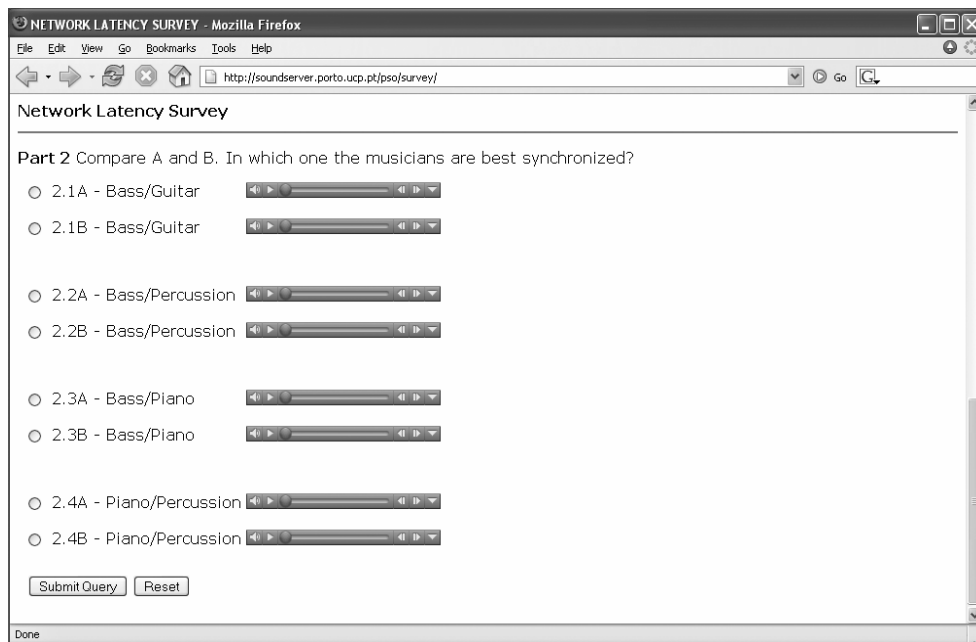


Figure 55. Online Survey for the evaluation of Individual Delayed Feed-Back Performances

<sup>52</sup> Continuation of the survey presented in Section 3.1.2 of this Chapter

Once again, the survey was submitted to 32 subjects with a dominant profile of music students from the School of Arts of the Portuguese Catholic University (53% with academic music training; 28% can play a musical instrument; 19% without any musical training). Song A always corresponded to the Normal Feed-Back Topology and Song B corresponded to an Individual Delayed Feed-Back Topology. Results are presented in the following Table:

Date	IP Address	Age	Gender	Education	Musical Training	2.1	2.2	2.3	2.4
27-04-2005 11:40	172.20.80.60	20	M	University degree	Academic training	B	B	B	B
20-04-2005 14:58	172.20.80.60	25	F	University degree	Academic training	B	B	B	B
19-04-2005 12:07	172.20.80.60	38	F	University degree	Academic training	B	B	B	B
29-06-2005 16:19	172.20.80.60	31	M	University degree	Academic training	B	B	B	B
18-04-2005 16:24	172.20.80.60	23	M	University degree	Academic training	B	B	B	B
18-04-2005 16:20	172.20.80.60	26	M	University degree	Academic training	B	B	B	B
02-05-2005 16:25	172.20.80.60	25	F	University degree	Academic training	B	B	B	B
02-05-2005 16:58	172.20.80.60	25	M	University degree	Academic training	B	B	B	B
03-05-2005 15:00	172.20.80.60	23	F	University degree	Academic training	B	B	B	B
03-05-2005 17:26	172.20.80.60	32	M	University degree	Academic training	B	B	A	A
06-05-2005 12:37	172.20.80.60	28	M	University degree	Academic training	B	B	B	B
07-05-2005 11:57	84.143.179.74	45	M	Post-doc	Academic training	B	B	B	B
11-05-2005 19:17	193.145.55.204	28	M	University degree	Academic training	B	B	B	B
11-05-2005 20:32	193.145.55.204	29	M	University degree	Academic training	B	B	B	B
12-05-2005 14:59	172.20.80.60	24	F	University degree	Academic training	B	B	B	B
19-05-2005 10:45	172.20.80.60	28	M	University degree	Academic training	B	B	B	A
24-05-2005 11:18	172.20.80.60	24	F	University degree	Academic training	B	B	B	B
28-04-2005 10:09	172.20.80.60	23	M	University degree	Can play musical instrument	B	B	A	B
28-04-2005 10:34	172.20.80.60	26	M	University degree	Can play musical instrument	B	B	B	A
28-04-2005 11:55	172.20.80.60	34	M	University degree	Can play musical instrument	B	B	B	B
03-05-2005 10:14	194.117.24.10	31	M	University degree	Can play musical instrument	B	B	B	B
03-05-2005 13:07	172.20.80.60	36	F	University degree	Can play musical instrument	B	B	B	B
07-05-2005 17:46	172.20.80.60	24	M	Secondary School	Can play musical instrument	B	B	A	B
11-05-2005 18:07	193.145.55.204	37	M	University degree	Can play musical instrument	B	B	B	B
11-05-2005 20:03	193.145.55.204	25	M	University degree	Can play musical instrument	B	B	B	B
16-05-2005 15:15	141.83.78.62	28	F	University degree	Can play musical instrument	B	B	B	B
23-04-2005 16:08	172.20.80.60	33	M	University degree	No Training	B	B	B	B
30-04-2005 17:54	192.35.246.5	36	M	University degree	No Training	A	B	A	A
02-05-2005 11:19	172.20.80.60	34	F	University degree	No Training	B	B	B	B
10-05-2005 11:23	193.145.56.194	26	F	University degree	No Training	B	B	B	B
A:	1	0	4	4					
B:	31	31	28	28					

Table 3. Results from the online Survey on evaluation of Individual Delayed Feed-Back

It is clear that in any case over 85% of subjects consider that IDF Topology (Song B) produces better results.

Based on this corroboration of the assumption that an IDF Topology allows better Individual Latency-Tolerance, the co-author of these experiments, Alexander Carôt, implemented a Delayed Feed-Back Feature in his Application for low-latency acoustic communication over the Internet entitled Soundjack (Carôt, A., 2004).

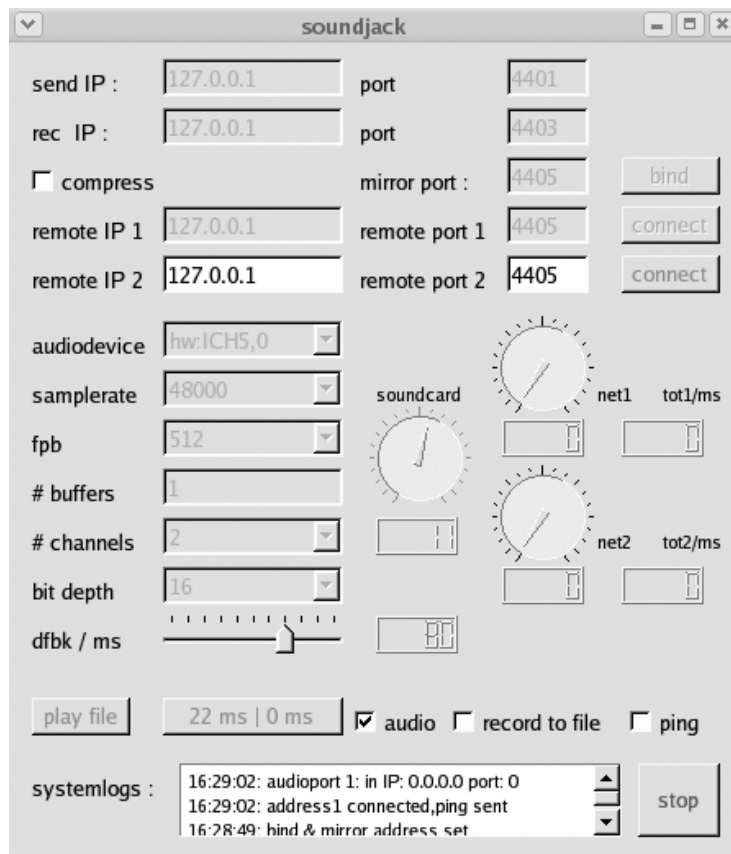


Figure 56. Soundjack Interface by Alexander Carôt

This application is based on the StreamBD research prototype software<sup>53</sup> from the CCRMA center at Stanford University. The interface allows users to manually match their Individual Feed-Back Delay with the Session’s Delay, by actuating on the slider “dfbk/ms”.

The concept of IDF was equally applied in the Public Sound Objects System discussed extensively in Chapter 5 of this dissertation.

<sup>53</sup> streambd is available from: <http://ccrma.stanford.edu/groups/soundwire/software/newstream/docs/>

## 4.2 Web services and Acoustic Applications

The World Wide Web International Consortium (W3C) for the development of Web standards, acknowledged since March 2001 the “W3C Web Services Activity” based on the following activity statement:

“Web services provide a standard means of interoperating between different software applications, running on a variety of platforms and/or frameworks. Web services are characterized by their great interoperability and extensibility as well as their machine-processable descriptions thanks to the use of Extensible Markup Language (XML), and they can then be combined in a loosely coupled way in order to achieve complex operations. Programs providing simple services can interact with each other in order to deliver sophisticated added-value services.” (Web Services Activity Statement from <http://www.w3c.org/>)

However, the idea of using the web to provide interoperability between remote software instances for Music Technology applications is prior the introduction of XML language in 1999 by the W3C.

The idea behind a remote graphical user interface and a processing engine that resides in a different host, taken to the most extreme situation on which a user can access the system from any place in the world using internet technology, is a concept similar to Sun’s “Nomadic Computing” concept (Gadol, S. and Clary, M., 1994).

Early applications were impelled by the benefit provided by Internet as a form of making specific software available to a broad spectrum of users. These software tools were typically dependent on special purpose hardware or proprietary experimental systems from companies or research groups.

Hence in 1995, with the support from Sun Microsystems, the Institut de Recherche et Coordination Acoustique (IRCAM) started a project of an on-line studio (Wöhrmann, R. and Ballet, G., 2002), based on client/server Web technology. The main purpose of this project was to provide access to some of IRCAM’s sound databases and sophisticated sound-processing tools like the phase vocoder SVP.

Access to this on-line studio was primarily conceived bearing in mind in-house access at IRCAM's intranet, since high speed network communication could be provided and it was not possible for each user to have an individual work-station with the required computing power for the studio applications.

A similar project began in 1997 with Ramon Loreiro (Loreiro, R. and Serra, X., 1997) at the Music Technology at Pompeu Fabra University in Barcelona (MTG), Spain, but with a slightly different scope.

The system provided a remote interface for a sound database and signal processing. Yet, it was primarily intended to be available in a simple and effective way for a broader community of users, granting access to cutting edge applications developed at the MTG.

With this project it was possible to have a web front end to Xavier Serra's Spectral Modeling Synthesis (SMS) (Serra, X., 1989) technique, based on Sound Modeling with sinusoids plus noise, which has diverse scientific and musical applications.

Research at MTG is highly focused on content base processing and analysis of musical signals, and therefore further projects based on the idea of providing added valued software services over the web have been developed. Namely, as part of the **Free-Sound Project**, presented in chapter 2, section 2.4.2.2, Jorn Lemon wrote the application **Mootcher** at the MTG during 2005. Mootcher is a Pure Data (PD) (Puckette, M., 1996a) (Puckette, M., 1996b) external which allows the access over the Internet to the Free-Sound Data-Base of Sound Samples.

With this application it is possible to embed the possibility to have remote access to an up-to-date global sample data-base in a PD patch. Mootcher also allows new and interesting ways to access these samples, such as: browsing the sounds in new ways using keywords, a "sounds-like" type of browsing; up and download sounds to and from the database, under the Creative Commons license; interact with fellow sound-artists.

At the MTG other projects were developed with the purpose of providing on-line web services related to the research carried out in the context of this dissertation.

### 4.2.1 Semiautomatic Ambiance Generation On-Line

The audio component of audiovisual productions has long been regarded as being of minor importance. Nevertheless, in the last years it is gaining interest for its evocative and overall

immersive experience of the audiences. Audio has an immense power, for creating the illusion of reality even when accompanying coarsely drawn cartoons.

Traditionally, from the film production process point of view, sound is broken into a series of layers: Dialog, Music and Sound Effects (SFX). SFX can be broken down further into Hard SFX—car doors opening and closing, and other foreground sound material. This SFX can be created through the process of Foley (made by humans, e.g.: footsteps from one side) or ambiance recording.

Ambiances also known as atmospheres are the background recordings of scenes and identify them aurally. They make the listener really feel like they are in places like an airport, a church, a subway station, or the jungle.

Considering this perspective the Semiautomatic Ambience generation On-Line project (Cano, P. and others, 2004) aims to provide remote web access to a database of SFX, which can be used to artificially create an atmosphere.

The application intends to provide an iterative process that departs from a semantic description of the desired atmospheres and that can be fine tuned manually over consecutive accesses to the data-base.

The system is based on a concept-based SFX search engine developed within the AudioClas project ([www.audioclas.org](http://www.audioclas.org)). The objectives of the project were to go beyond current professional SFX provider information retrieval model, based on keyword-matching, mainly through two approaches:

- Semantically-enhanced management of SFX using a general ontology, WordNet (Miller, G. A., 1995);
- Content-based audio technologies which allow automatic generation of perceptual meta-data (such as prominent pitch, dynamics, beat, noisiness).

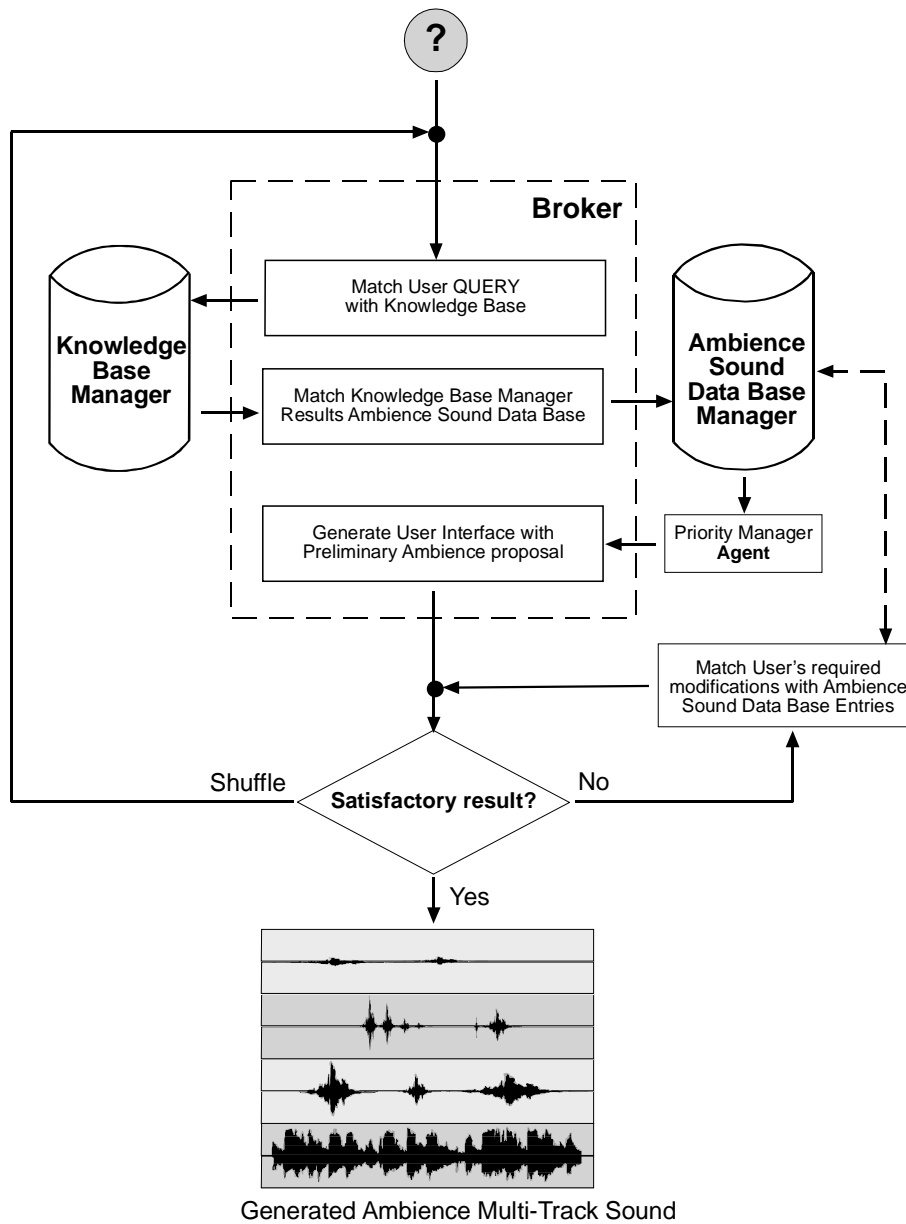


Figure 57. Flow diagram of the Semiautomatic Ambience generation On-Line system.

The final result is delivered to the client application in the form of multi-track sound file, ready to be imported to a non-linear audio editing system.

#### 4.2.2 Data Sonification on Demand

The development of a proof-of-concept application in the context of this dissertation, **Public Sound Objects** (PSOs), followed key requirements discussed on Chapter 5.

One of the main requirements was system scalability, so that, its architecture could be adapted to diverse on-line applications other than a Shared Sonic Environment. This was the case in the **Sound Data Mining** (SDM) Project developed at the Research Center for Science and Technology of the Arts (CITAR) at the Portuguese Catholic University.

The work developed in the scope of this project consists in the application of spatial data mining techniques to various fields. Sound is used as one of the means for information presentation and data exploration. Spatial databases hold information on geo-referenced data, i.e., data regarding the location and shape of geographic features. Spatial data includes both topological and geometric data.

As with other types of large databases, one of the most important and difficult aspects of spatial databases is the extraction of knowledge. Spatial databases typically have huge amounts of spatial data that render the human ability to analyze useless, making it necessary to use automatic methods of analysis and knowledge discovery, or extraction.

As defined in (Koperski, K. and Han, J., 1995), **spatial data mining** is the extraction of implicit knowledge, spatial relations, or other patterns not explicitly stored in spatial databases. Spatial data mining techniques enable us to obtain information that would be difficult to get otherwise.

A simple example of information that could be obtained from a spatial data mining process is the correlation of a cholera outbreak in London in the 19th century, to a contaminated water pump located on Broad Street. Although this correlation was done “manually” in 1854 by Dr. John Snow (Snow, J., 1855), we might imagine an extrapolation to nation or world-wide disease analysis that would demand for automatic processing.

**Sonification** is the use of non-speech audio to convey information (kramer, G. and others, 1999). It is of special interest when there is a high data volume and number of variables; in these cases it may be useful to present a part of the information visually and a part through audio.

Audio can be used to increase perception of the information that is being graphically displayed, or it may be used to present information that is not displayed visually. The output of a spatial data mining process can take many forms, e.g., clustering, classification, prediction, etc. The use of acoustic information becomes more important as the graphical capacity of the user interface diminishes. This is especially true in the case of mobile devices where the graphical display is very limited, not only in terms of size, but also in color depth and resolution.



Sonification techniques become especially interesting when the client application runs on these graphically limited devices such as mobile phones or PDAs (Personal Digital Assistants).

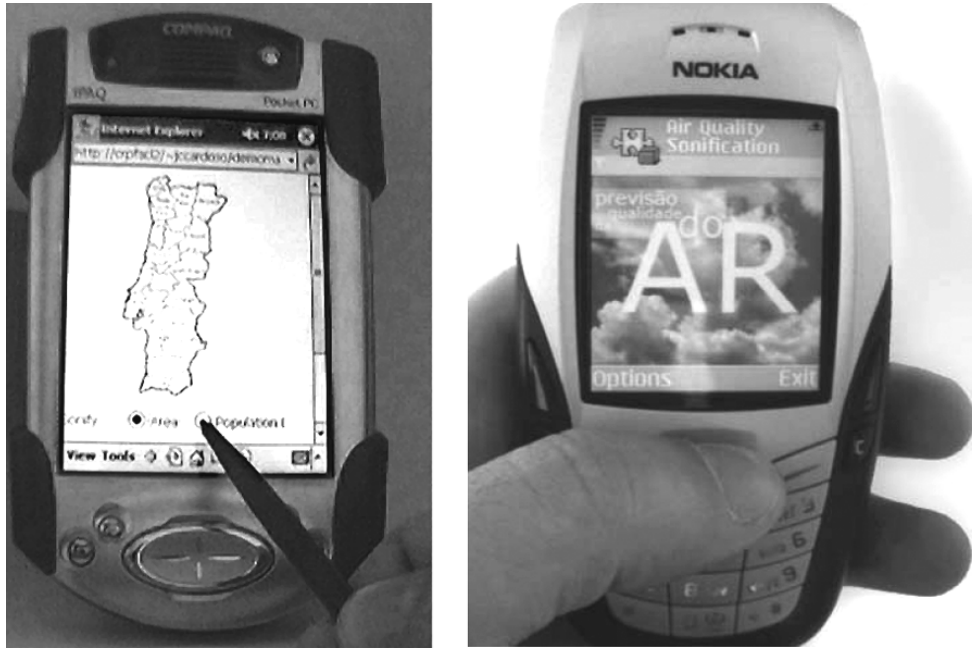


Figure 58. Interfaces of the SDM Project for mobile devices

Furthermore, remote access to a powerful and flexible sound synthesizer enables computation which cannot be performed in these clients.

In the paper “Soundserver: Data Sonification On-Demand for Computational Instances” (Cardoso, J. and others, 2004), co-authored by the author of this thesis, an architecture was presented for a sonification client-server system based on an audio synthesis engine similar to the PSOs System.

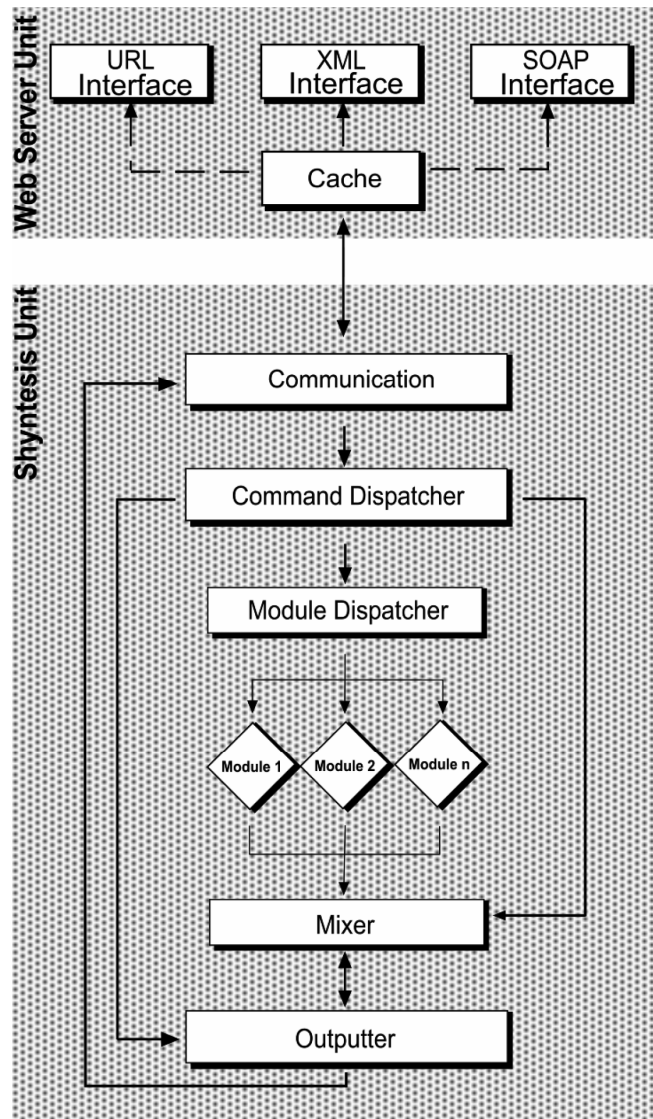


Figure 59. Architecture of the SDM System

A centralized server unburdens clients of little audio synthesis capabilities from the burden of synthesizing the sounds, maintaining the possibility of geographical displacement of the clients.

### 4.3 Chapter Conclusions

The Acoustics of a performative space determines to a great extent the form and style of a musical performance. In this chapter is presented an analysis to perception issues related with the use of internet audio connectivity as an acoustic media to perform music. Even though here are different technical and operational issues which cause disruptive effects in s geographically

displaced performance, the main focus of this chapter is on network Latency. Based on an experimental study of latency tolerance to music performance it is proposed the concepts of **Latency Adaptive Tempo and Dynamics** (LAT and LAD) and **Individual Delay Feed-Back** (IDF). These ideas are also introduced in chapter 5 as software functions implemented in the Public Sound Objects Systems.

In this chapter is also briefly presented further examples of internet acoustic applications, which have been developed based on preliminary work and concepts derived from the study developed during this doctorate research, both in terms of network architectures and models presented in Chapter 3 or communication schemes implemented taking into account the results of Network Latency Studies. In fact, the synthesis engine and overall system architecture of the Sound Data Mining system presented in section 4.2.2 derives from a preliminary version of the Public Sound Objects System presented in the following chapter.

# Chapter 5

## The Public Sound Objects: A System Prototype for Experimental Research

In his book “Weaving the Web” the creator of the World Wide Web(WWW), Tim Berners Lee, explains his dream for the future:

“I have a Dream for the Web...and it has two parts.

In the first part, the web becomes a much more powerful means of collaboration between people. I have always imagined the information space as something to which everyone has immediate and intuitive access, and not just to browse, but to create.

(...) In the second part of the dream, collaborations extend to computers. Machines become capable of analysing all the data on the Web – the contents, links and transactions between people and computer:. A ‘Semantic Web’ which should make this possible (...)” (Berners-Lee, T., 2000)

The World Wide Web is a simple and effective paradigm to retrieve and share information. The notion of hyperlink had a tremendous influence in the Internet’s expansion, and it preceded present developments on internet collaboration in both mentioned approaches (human to human, and computer to human collaboration).

In fact, today web browsers are used as a sort of multiplatform virtual machine, which embeds multimedia applications written in advanced programming environments, such as Java, Flash or Shockwave.

Today, the WWW is a default front-end to use Internet resources, such as e-mail, news feeds or even FTP file transfers, which can now be performed within a web browser by a growing majority of Internet users

Considering the results from the survey on Computer-Supported Cooperative Work for Music Applications presented in Chapter 2, it is clear that one of the most promising approaches to Networked Music practices relies on Shared Sonic Environments (SVEs). This class of applications directly explores distinctive characteristics of Internet computing, i.e., its distributed and shared nature.

SVEs are oriented towards the general and anonymous internet user and therefore WWW provides the most appropriate communication layer for such systems. This way, the target audience is quite large and is guaranteed to have basic common knowledge of interaction with the WWW front-ends.

The **Public Sound Objects** (PSOs) project consists of the development of a networked musical system, which is an experimental framework to implement and test new concepts for on-line music communication. It not only serves musical purpose, but it also facilitates a straightforward analysis of collective creation and the implications of remote communication in this process.

The PSOs project approaches the idea of collaborative musical performances over the Internet as a **Shared Sonic Environment** aiming to go beyond the concept of simply using computer networks as a channel to connect performing spaces. It runs entirely over WWW, and its underlying communication protocol (Hypertext Transfer Protocol - HTTP<sup>54</sup>), in order to achieve the sense of a **Public Acoustic Space** where anonymous users can meet and be found performing in collective Sonic Art pieces.

The system itself is an interface-decoupled Musical Instrument, in which a remote user interface and a sound processing engine reside with different hosts in an extreme scenario where a user can access the synthesizer from any place in the world using a web browser.

Specific software features were implemented in order to reduce the disruptive effects of network latency, such as, dynamic adaptation of the musical tempo to communication latency measured

---

<sup>54</sup> More information about http is available from: <http://www.w3.org/Protocols/>

in real-time and consistent sound panning with the object's behavior at the graphical user interface.

## 5.1 Community Music and Sound Objects

Community-driven creation, results in a holistic process, i.e., its properties cannot be determined or explained by the sum of its components alone (Smuts, J., 1926). A community of users involved in a creation process, through a Shared Sonic Environment, constitutes a Whole in Holistic sense.

According to Jan Smuts (1870-1950), the father of Holism Theory, the concept of a Whole implies its individual parts to be flexible and adjustable. It must be possible for the part to be different in the whole from what it is outside the whole. In different wholes a part must be different in each case from what it is in its separate state.

Furthermore, the whole must itself be an active factor or influence among individual parts, otherwise it is impossible to understand how the unity of a new pattern arises from its elements. Whole and parts mutually and reciprocally influence and modify each other.

Similarly, when questioning object's behaviors in Physics it is often by looking for simple rules that it is possible to find the answers. Once found, these rules can often be scaled to describe and simulate the behavior of large systems in the Real World.

This notion applies to the Acoustic Domains through the definition of **Sound Objects** as a relevant element of the music creation process by Pierre Schaeffer in the 1960's. According to Schaeffer, a Sound Object is defined as:

“Any sound phenomenon or event perceived as a coherent whole (...) regardless of its source or meaning” (Schaeffer, P., 1966).

Sound Object (*I'object sonore*), refers to an acoustical object for human perception and not a mathematical or electroacoustical object for synthesis. One can consider a sound object the smallest self-contained particle of a Soundscape (Schafer, M., 1977).

Defining a universe of sound events by subsets of Sound Objects is a promising approach for content-processing and transmission of audio (Amatriain, X. and Herrera, P., 2002), and from a psychoacoustic and perceptual point of view it provides a very powerful paradigm to sculpt the

symbolic value conveyed by a Soundscape. Adding metaphorical value to a Sound Objects enhances the awareness of individual Sound Objects within a Soundscape.

On the other hand, the symbolic value of the Sound Object might also change depending on the context in which it is presented. In many applications such as Auditory Users Interfaces, Sound Objects must be simple and straightforward, so that there is no ambiguous understanding of what they intend to represent.

However, in an artistic context the scope for the user's personal interpretation is wider. Therefore such Sound Objects can have a much deeper symbolic value and represent more complex metaphors. Often there is no symbolic value in a sound, but once there is a variation in one of its parameters it might then convey a symbolic value.

A typical example is the use of white noise to synthesize wind sound. If we listen to continuous white noise it might not represent a very strong metaphor, although we could relate it with some meaning depending on its context. It can for instance be perceived as an offline transmission device. However, if we apply a band pass filter to this sound varying its central frequency, even out of any special context, we can perceive the result as the very familiar natural sound of wind blowing.

All these ideas about Sound Objects and the Holistic nature of community music are the basis for the main concept behind the Public Sound Objects System. In fact, in PSOs raw material provided for each user, to create his contribution to a shared musical piece, is a simple Sound Object. These Sound Objects, individually controlled, become part of a complex collective system in which several users can improvise simultaneously and concurrently.

In the system a server-side real-time sound synthesis engine provides an interface to transform various parameters of a Sound Object, which enables users to add symbolic meaning to their performance. Musically the outcome of PSOs performance relates to idea of Musical Sound presented by Douglas Kahn:

“The line between sound and musical sound stood at the center of the existence of avant-garde music, supplying a heraldic moment of transgression and its artistic raw material, a border that had to be crossed to bring back unexploited resources, restore the coffers of musical materiality, and rejuvenate western art music” (Kahn, D., 1999).

## 5.2 System Overview and Architecture

A preliminary specification of the Public Sound Objects system was published in (Barbosa, A. and Kaltenbrunner, M., 2002), and the first prototype was implemented in December 2002. The system is publicly available on-line from the **URL: <http://www.iaa.upf.es/~abarbosa/>**.

The overall system architecture was designed with respect to the following key factors:

1. It is based on a Centralized Server Topology supporting multiple users connected simultaneously and communicating amongst themselves through sound;
2. It is a permanent public event with special characteristics appealing both to a “real world” audience and to an on-line virtual audience, since an on-site installation version of the system resides at the server’s physical location;
3. The user interface and the sound synthesis engine offer a constrained sonic creation paradigm, which provides coherence amongst individual contributions;
4. Even though the user interface does not resemble any traditional musical instrument, the parameters of sound which a user can control are the usual ones (Tempo, Pitch, Dynamics and Timbre)
5. The system is scalable and modular allowing expansion to diverse setups.

The PSOs system is based on client-server architecture. Clients control a visual interactive interface, while the server controls all computation regarding the sound synthesis and transformation and all features for a local installation. It is an extreme example of an Interface-Decoupled application where the synthesis engine is separated from the user interface over a Large Area Network (Barbosa, A., Kaltenbrunner, M. and Geiger, G., 2003).



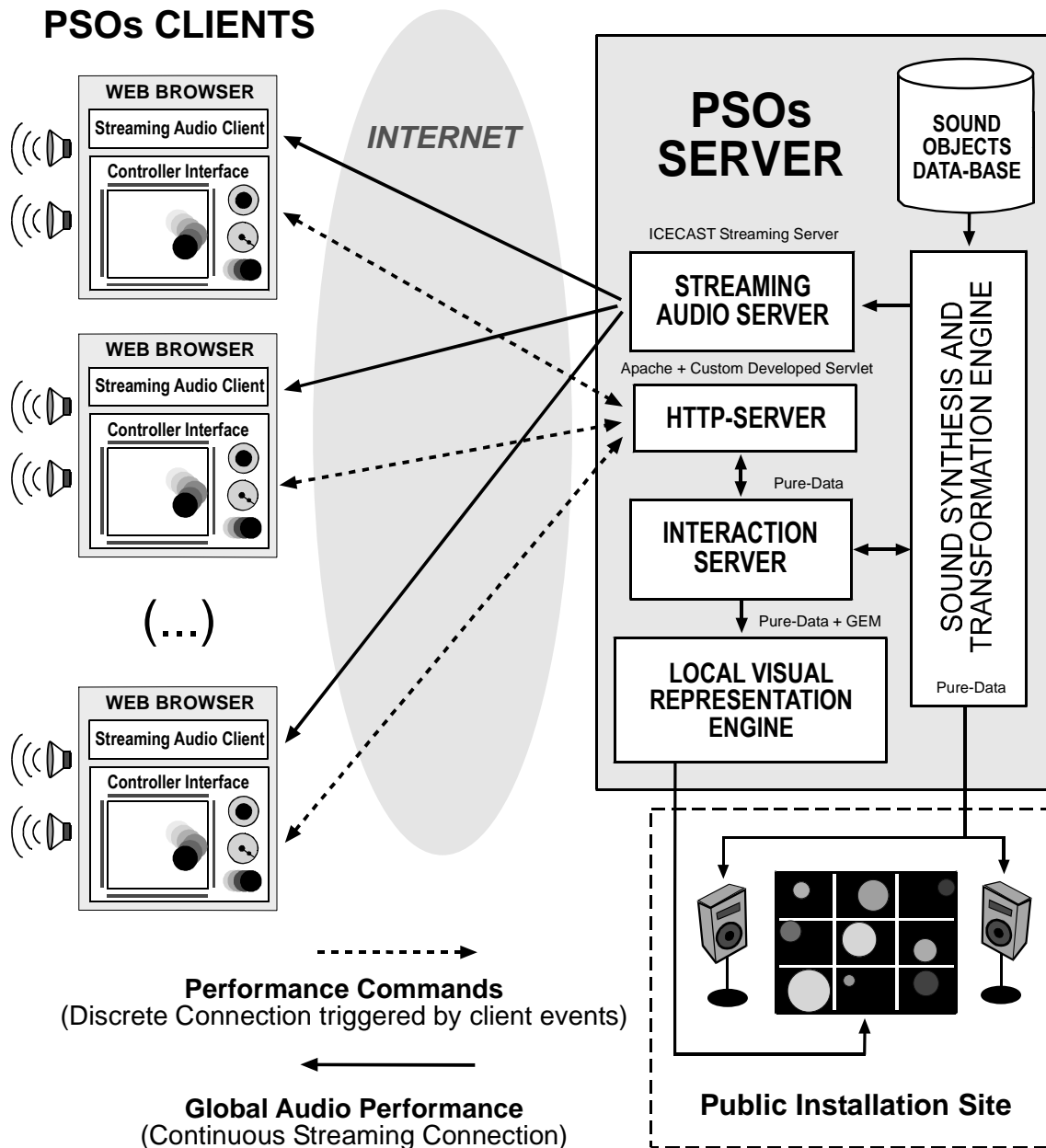


Figure 60. The Public Sound Objects Architecture

Clients communicate with the server through HTTP by sending and receiving packets of data. There are several types of data packets that the clients can send but the most important ones are the ImpactPacket – which informs the server that the bouncing ball has hit one of the walls; the ControlPacket – which tells the server that the user has changed the value of one of the interface controls; and the PingPacket – which is used to measure the network delay between the client and the server.

The server packets are received by a Web application that reroutes them to the Interaction Server – a module of the PSOs Server that manages clients, instruments and the events generated by the PSOs Client. Depending upon the type of data packet received, a sound can be generated by the Synthesis and Transformation Engine and then streamed back to the client by the Streaming Audio Server, or the visual representation of the client can be updated at the installation site by the Local Visual Representation Engine, or both.

Server and Clients have different modules:

### 5.2.1 Web Server

Clients connect to the PSOs Server through standard Hypertext transport Protocol HTTP connections. Although the initial choice was to implement UDP based communications – faster than a TCP based protocol like HTTP – the idea had to be abandoned for two main reasons:

- Most firewalls block all unknown UDP traffic which meant that a great number of users would not be able to access our server. Also, increasing the difficulty of deploying the PSOs Server for the same reasons: UDP traffic would have to be allowed at a specific port by the firewall.
- Some browsers' security policies for Java applets only allow them to make connections using the HTTP protocol.

In order to overcome these restrictions a communication system was realized using a “firewall generally allow” protocol: HTTP. For this a server application was implemented, using the Java Servlet technology, which acts as a proxy between the PSOs Client applet and the Interaction Server. Basically, this servlet just passes data received from the PSOs Client to the Interaction Server and vice-versa.

### 5.2.2 Communication Layer

The Interaction Server is a central piece in the PSOs Server. It's a Pure Data (PD) module that receives data packets in the form of UDP datagrams from the clients (through the HTTP Server) and acts accordingly to the type of packet received.

A custom PD object had to be implemented for the reception of the UDP datagrams – which was called Extended Netreceive [xnetreceive] – since existing objects for this purpose don't allow PD to acquire the IP address and port number of the client that initiated the communication.

The packet types defined so far are as follows:

- **AvailableInstruments:** When the Interaction Server receives this type of packet it sends as response the numbers of the instruments that are available. Instruments were numbered 1 to 9.
- **LockInstrument:** This type of packet is sent to the server when the user chooses one instrument to play. The instrument number is specified in the packet. The Interaction Server will check that the instrument is still available and will respond with a True/False result depending on whether the instrument was successfully locked or not. When an instrument is locked it can only be used by the client that locked it.
- **UnlockInstrument:** Informs the server that the user is done with the instrument specified by the instrument number in the data packet. The Interaction Server will unlock the instrument, which will then become available to other clients.
- **ImpactPacket:** This is the most used packet. It tells the server that the bouncing ball has hit a wall and that a sound should be generated. The Interaction Server passes these packets along to the Synthesis and Transformation to the Local Visual Representation Engines. Among other information, these packets specify the instrument number, the value of the wall sliders, the speed of the ball, the ball's size, the wall that was hit and what point of the wall was hit and the size of the ball's trail. This information is then used by the Synthesis and Transformation Engine to generate a sound accordingly to the parameters set by the user in the PSOs Client interface. It is also used by the Local Visual Representation Engine to update the visual representation of that user.
- **ControlPacket:** This type of packet is of interest only to the Local Visual Representation Engine. The information that is sent is the same as the ImpactPacket but the events that trigger transmission are different. ControlPackets are sent whenever the user changes the speed, size or trail size of the bouncing ball. The Interaction Server passes these packets along to the Local Visual Representation Engine so that the installation site can be updated.

- **PingPacket:** These packets hold no direct information; their sole purpose is to allow the PSOs Client to determine the network delay between the client and the server. The Interaction Server merely sends back an empty reply to the client.

The other main task of the Interaction Server is to manage the connected PSOs Clients. If a client gets disconnected from the network without having sent an unlock packet, the instrument currently locked by that client would never again be available.

It is the job of the Interaction Server to detect such situations and to automatically free the instrument up. This is done with timeouts, i.e., if a client remains more than a fixed amount of time without contacting the server that client is removed from the list of currently connected clients and its instrument released.

### 5.2.3 Synthesis and Transformation Engine

The Synthesis and Transformation Engine is responsible for the sound generation in response to the PSOs Clients' generated events.

The main advantage of separating the controller from the synthesizer is that this way the synthesis and transformation engine can be implemented in a computer with much higher requirements of digital sound processing performance and efficiency, than it could be demanded from general client computers.

The synthesis engine is a PD patch automatically loaded by the Interaction Server. It receives ImpactPackets from the Interaction Server (PD lists) and generates a sound according to the values specified therein.

The parameters taken from these data packets are actually passed on to one of nine synthesis modules.

At this time, the engine has nine modules that correspond to the nine instruments available to users. Since each module is different and independent, the same parameter can have a different meaning for different modules. These modules are:

- **Karplus-Strong Guitar:** As the name suggests, this is an implementation of the Karplus-Strong algorithm for a plucked string sound implemented in PD (Karplus, K. and Strong, A., 1983).

- **FM Synthesizer:** A frequency modulation synthesizer (Chowning, J., 1973).
- **Modal Impact Vibraphone:** An attempt to produce vibraphone-like sounds using Modal Impact physical models implemented for PD, available from the Sounding Objects Project (Rocheso, D. and Fontana, F., 2003).
- **Piano, Percussion, Violin, Orchestra, Tabla and Poet Samplers:** These are in fact only one module, loaded with six different sounds. The sampler was implemented in PD and used six voices, which proved to be enough not to overload the system, for the worst case scenarios (nine users connected with high-tempo performances).

The sound generated by these modules is streamed in MP3 format, using the [shoutcast~] PD object to an audio streaming server. The streaming server is Icecast2 for Windows. The MP3 Stream between PD and the Streaming server is set to a fixed bit rate to avoid more jitter in the latency (variable delay).

Each user can choose one of these modules as the sound generating engine from the PSOs entry screen at the client instance.

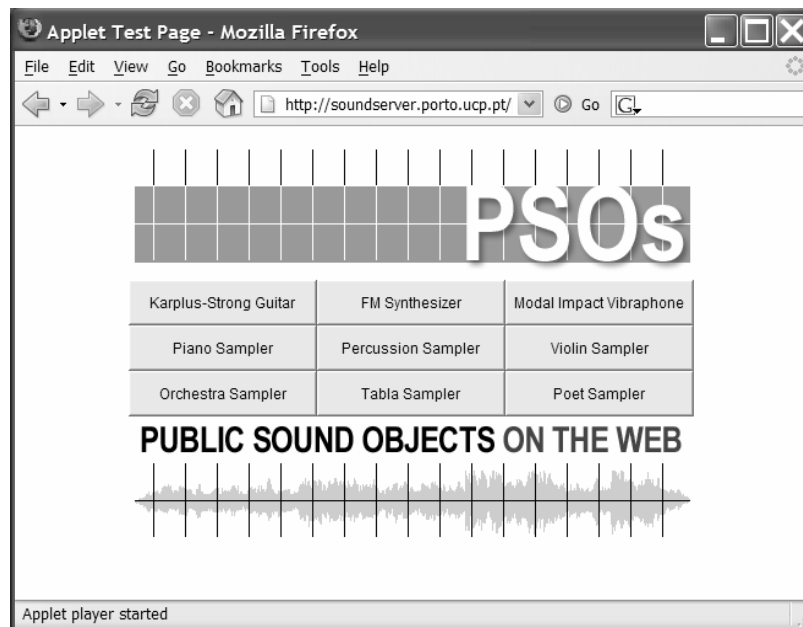


Figure 61. Entry Screen for PSO Client Version 3

Upon loading the entry web page, if a Sound Module is taken by another user, its button on screen will be off and the Module will only be available when it is released.

Once the Sound generating engine (instrument) has been selected, the web-browser loads the controller interface applet, which connects to the interaction server, registers and initializes a user session.

## 5.3 User Interface

The graphical user interface (GUI) implementation was developed along the following requirements:

1. It should enable the user to contribute to the ongoing musical performance by transforming the characteristics of a visual Sound Object representation, sending normalized parameters (control data) to the synthesis engine over the network;
2. It should embed a streaming audio client to receive the global Soundscape.
3. It should run on a web browser.
4. The interface application should be able to allow manipulation of each of the modifiers' parameters in the synthesis engine; It should also be articulated with the installation site setup;
5. The GUI itself should be a **behavior-driven** metaphorical interface, avoiding a flat mapping of parameters, such as faders or knobs; providing automatic periodical behavior for the Graphic Objects as a sound controller allows larger timescales in the user action, which tends to be more appropriate for a system with delayed acoustic feedback.
6. The user interface should not resemble any traditional musical instrument; however, the controllable sound parameters should be based on Tempo, Pitch, Dynamics and Timbre (most familiar to a generic user).

Considering these requirements it was considered that a good approach for the GUI could be based on a bi-dimensional graphical metaphor of an ever-going bouncing ball enclosed in a square shape box.

This idea was inspired by the work of Kiyoshi Furukawa, Masaki Fujihata and Wolfgang Muench developed in 1999 at the ZKM centre for Art and Media in Karlsruhe, Germany. The CD-Rom “Small Fish” (Furukawa, K., Fujihata, M. and Muench, W., 1999).

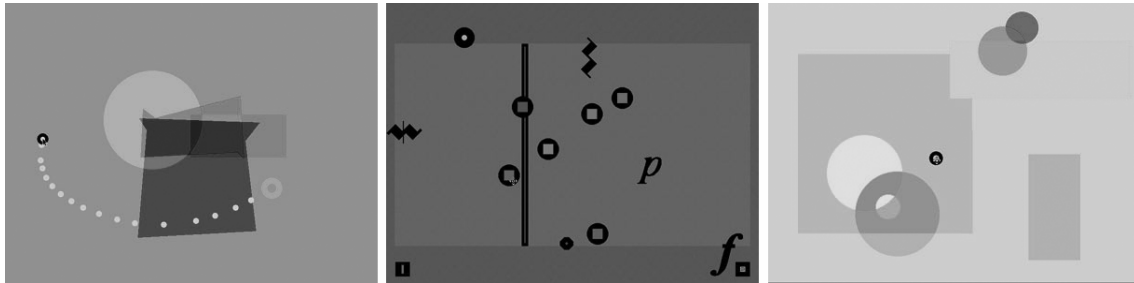


Figure 62. Small Fish Visuals by Masaki Fujihata

Transposing the Small Fish bouncing balls generative music paradigm to a Network Music setup and geographically separating the user interface from the sound synthesis engine correspond to the requirements of PSOs concept because:

1. The interface is behavior driven;
2. Taken to most simple case a bouncing ball corresponds to one visual object that can represent a Sound Object;
3. It is simple to use by non practicing musicians;
4. It is simple to relate visually with the effects of Network latency, since there will be a time lag between the moment when a ball hits a wall and the correspondent sound reproduced at the client software;
5. It can provide control over the most typical musical parameters (Tempo, Pitch, Dynamics and Timbre).

Hence, the first prototype of the interface was written in 2001 using Macromedia Flash and piano samples as a draft for the Sound Synthesizer.

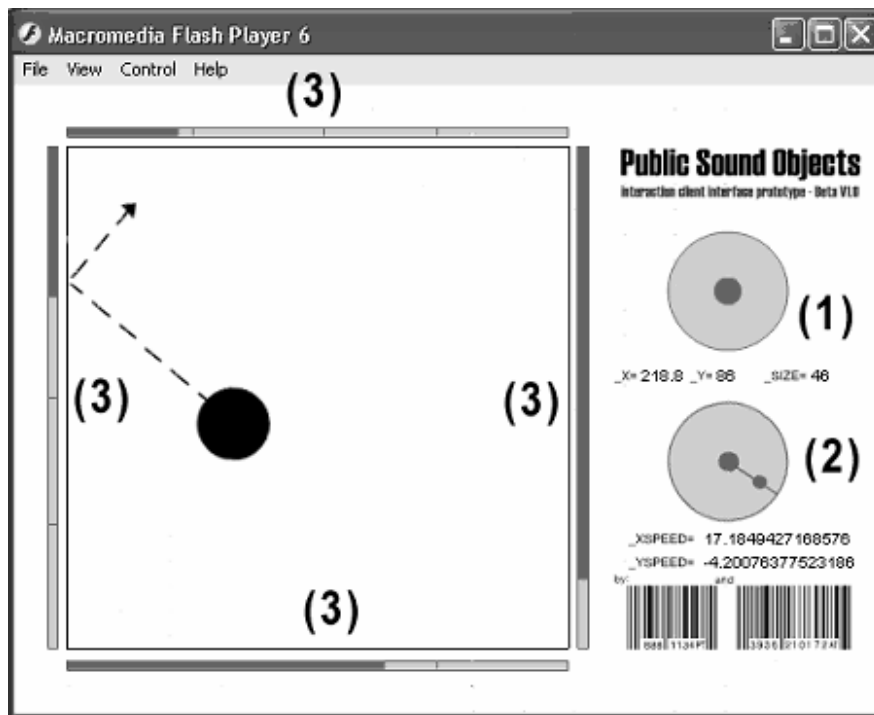


Figure 63. First prototype of the PSOs interface developed in Flash

When the ball hits one of the walls a corresponding Sound Object is triggered. The ball moves continuously and the user can manipulate its size, speed and direction by actuating on:

- (1) **The Ball size:** This control is directly proportional to the Bouncing Ball size. The ball size is one of the parameters that influence Musical Tempo.
- (2) **The Ball speed and Direction:** This control allows changes in a (x,y) vector, setting simultaneously the ball's speed and direction. These are the main parameters that influence Musical Tempo.
- (3) **The Pitch of Triggered Sounds:** Contiguous to each wall there is a fader, which allows to modify the pitch of a sound triggered by the ball's impact. The four walls have independent pitch, allowing the creation of melodic and rhythmic sound structures.

Departing from the Flash Mockup a similar interface was written in Java providing the necessary coupling with the synthesizer through the internet, as described in section 5.2.3 of this chapter.

Java language was chosen for the following reasons:



- It's a High-level object-oriented language
- It is Free
- It runs in multiple platforms (Java Virtual Machine)
- It has web deployment strategies (Applets, Server-side Java Applications, Servlets, etc.)
- Trough Java sockets programming, applications can access end-to-end transport protocol, namely the Transport Control Protocol (TCP) and User Datagram Protocol (UDP) in TCP/IP networks, without bothering with underneath network technology (e.g., Ethernet, Token Ring, ATM, etc)<sup>55</sup>.

The first implementation of a working system was realized in 2002 and presented in (Barbosa, A. and Kaltenbrunner, M., 2002).

In this version, when the ball hits one of the walls a network message is sent to the central server where the corresponding Sound Object is triggered, streamed back to the user in a stereo mix of all the sounds being triggered at the moment. In the web browser the streaming client is embedded in a separate frame from the GUI.

---

<sup>55</sup> Network protocols were discussed in chapter 4, section 4.1.1

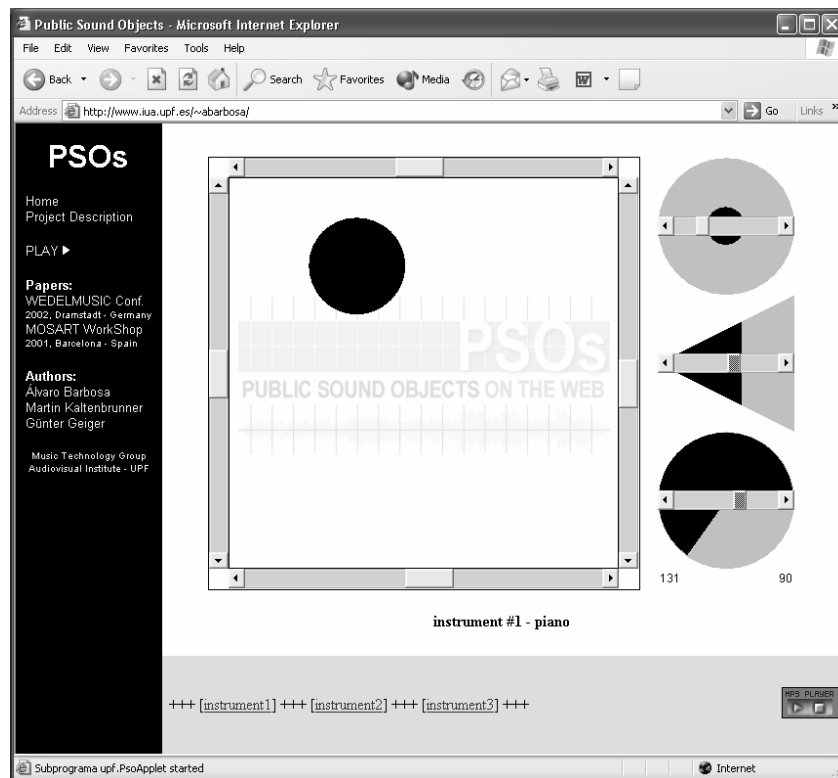


Figure 64. First Java implementation of the PSOs interface.

In the period of 2003-2004 the system was extended and refined to meet new requirements

- It was intended to perform in different host devices (PDAs, Touch-Screens and Web Banners);
- It should match the Sound Synthesis Engine developments (described in 5.2.3);
- It should support a physical installation at the servers Site.
- It should incorporate distinctive software features that resulted from recent research conducted in the context of this dissertation.

The new GUI had an entry screen, which allowed users to chose one of nine possible Sound Objects to manipulate, and only then their session would be registered and initialized on the server. The new control interface had a new control feature, the ball's tail extension that corresponds to the number of replicas of the delay applied to the Sound Object when its triggered at the server.

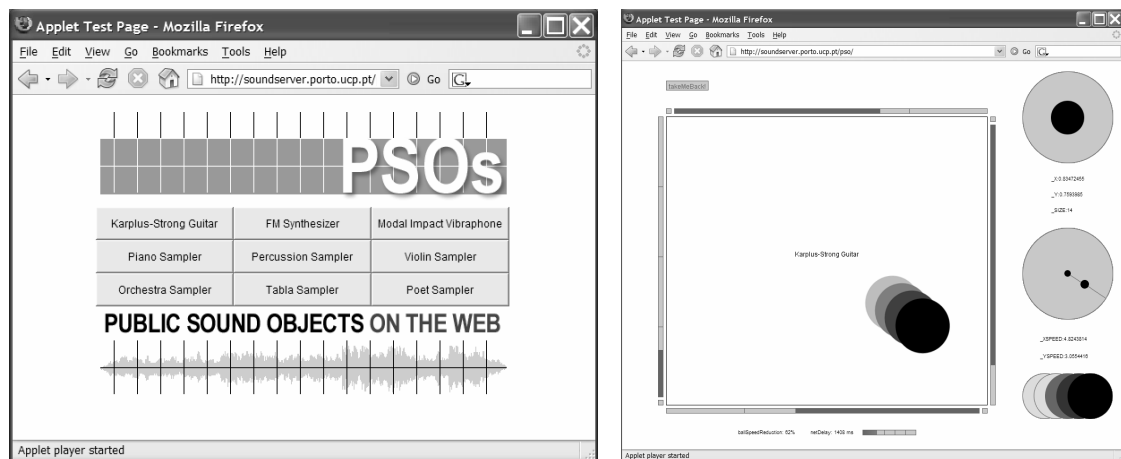


Figure 65. PSOs GUI version released in 2004; Entry screen for nine sound objects and controller with ball tail and real-time network latency measurement.

This version of the software already included a real-time network latency measurement. Further developments of the system were oriented towards the previously mentioned requirements regarding a multi-platform implementation and the installation site.

### 5.3.1 Multi-Platform Implementation

Initially, the PSOs GUI was intended to run in three diverse interfaces, A common Desktop Computer (the main application), a **Desktop Computer** equipped with a 14'' **Touch-Screen** and a Personal Digital Assistant (**PDA**) I-Pack Pocket PC from the manufacturer Compaq, running the Windows CE operating System. The interface had to be designed for each platform (in the Touch screen faders were wider to facilitate touch, and in the PDA controls were reduced to the minimum, since the screen resolution is very small), however only the PDA had specific hardware specificities in what concerned computation.

The Java programming environment used to write the application was J2SDK 1.4, however, in order to maintain compatibility with the PDA, all code was written in order to be compilable in Java version 1.1 (supported by Windows CE)



Figure 66. Desktop, Touch Screen and PDA Interfaces for PSOs

To generate graphics at the applet interface, the API AWT was used (Abstract Window Toolkit), once again to facilitate the integration in Windows CE, since this is the basic graphic API for Java and therefore it is “lighter”.

However, the software was written in such a way that it detects the hardware capabilities in runtime, and for the Desktop platforms it allows the use of Graphics2D Java Class, which allows graphics with anti-aliasing. Graphics for the PDA do not support anti-aliasing, since this class is not supported.

For the Audio Streaming reception a third party application is used, the jGui player Applet<sup>56</sup>, a streaming client for MP3<sup>57</sup> and Ogg Vorbis<sup>58</sup> audio formats. jGui requires Java 1.3, and it is not embedded on the controller applet. It is a separate instance, in the web page and it is transparent to the user, since its size is set to zero.

A final development of the PSOs GUI was released in June 2005, which approaches the ideas of involuntary exposure of this system to a general audience, by the means of interactive Banners.

**PSO Banner** version is a simplified version of the interface (almost as the PDA version, but with a banner aspect ratio), that can easily be embedded in a web page as any publicity banner.

---

<sup>56</sup> jGui is available from: <http://www.javazoom.net/applets/jlguiapplet/jlguiapplet.html>

<sup>57</sup> More information about MP3 format is available from the MPEG 1 – Layer 3 webpage: <http://www.chiariglione.org/mpeg/standards/mpeg-1/mpeg-1.htm>

<sup>58</sup> More information about Ogg Vorbis format is available from: <http://www.vorbis.com/>

**Interactive Banner Music** is an approach to be further explored for massive participation in collaborative Shared Sonic Environments.

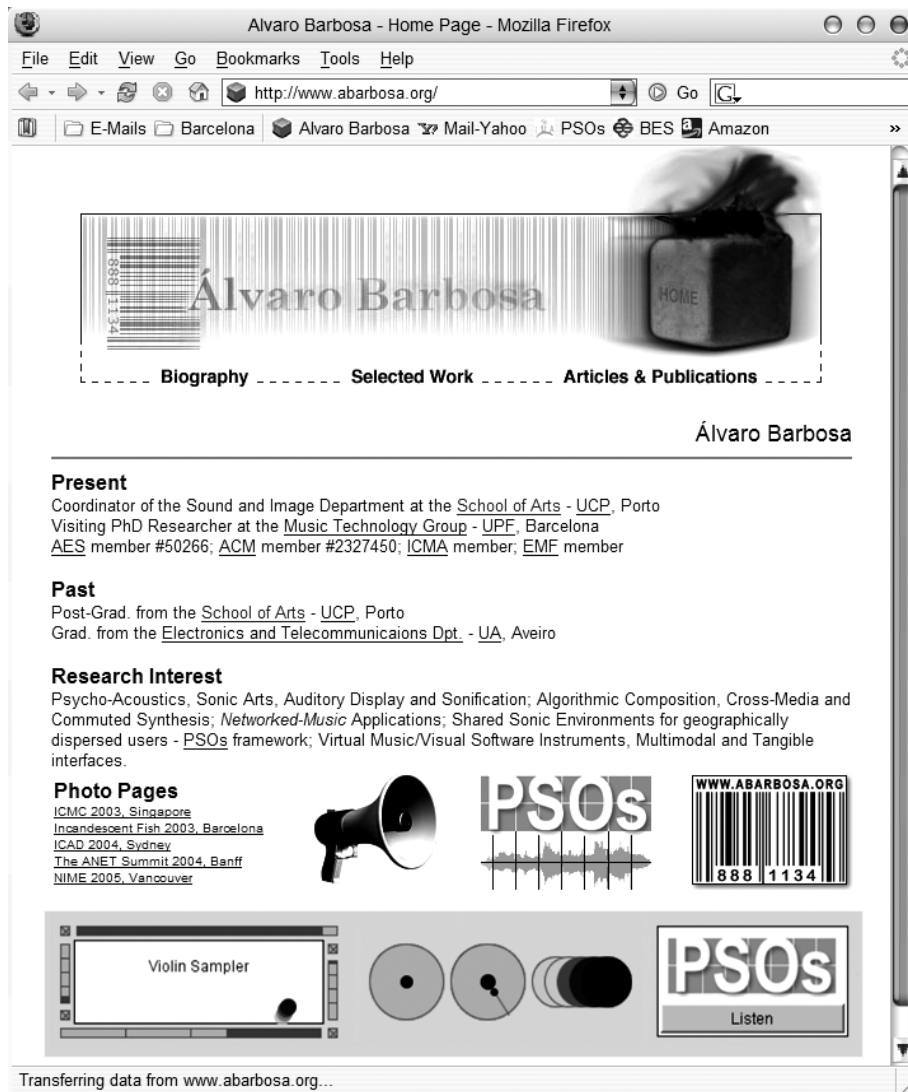


Figure 67. PSO Banner embedded at the Home Page for <http://www.abarbosa.org/>

The html code to embed a PSO banner Applet, including the streaming client, is as follows (it only contains absolute references):

```
<applet codebase="http://soundserver.porto.ucp.pt/psos" name="PSOApplet"
code="PSOApplet" archive="psos-banner.jar" width="630" height="100"
align="middle">
<param name="port" value="80">
<param name="minddelay" value="0">
```

```
</applet>
<applet codebase="http://soundserver.porto.ucp.pt/pso" name="player" width="0"
height="0" code ="javazoom.jlgui.player.amp.PlayerApplet"
archive="lib/jlguiapplet2.3.jar,lib/jlgui2.3-
light.jar,lib/tritonus_share.jar,lib/basicplayer2.3.jar,lib/mp3spi1.9.1.jar,lib/jl0.4.jar,lib/c
ommons-logging-api.jar">
<param name="scriptable" value="false">
<param name = "skin" value ="skins/bao.wsz">
<param name = "start" value = "yes">
<param name = "song" value ="http://soundserver.porto.ucp.pt:8000/pso.mp3">
<param name = "init" value ="jlgui.ini">
<param name = "location" value ="url">
<param name = "useragent" value ="winampMPEG/2.7">
</applet>
```

- Upon loading a Sound Object is randomly assigned (from the ones available at that moment) and the performance starts immediately. By clicking on the “Listen” button, the user can chose not to control any Sound Object and simply listen to the sonic performance of connected users at that moment.

### 5.3.2 Installation Site

Since the first specifications of the project it was intended that at some point an additional feature of the system should be developed, consisting of a physical installation located at the server site that could receive presential visitors. These visitors should be able to interact in a local client with geographically displaced virtual participants in the Shared Sonic Environment. By the fact that high-end technology is available at the server site, a variety of different interfaces should be implemented (Desktop, PDA and Touch-Screen) and a global visual representation of connected users would provide a more immersive experience for presential visitors about the paradigm they have just joined.

The implementation of PSOs Installation was realized in October 2004 at the Porto School of the Arts, Portugal, Where the main PSO server is hosted (<http://sounserver.Porto.ucp.pt/pso/>).



Figure 68. Mockup design of PSO Installation and the real implementation at Porto School of the Arts, October 2004

The Local Visual Representation Engine outputs the visual representation of the bouncing ball model of all the connected PSOs Clients, at the server's physical location. It consists of a PD patch that uses the Graphics Environment for Multimedia (GEM) external for graphics output, using information from ImpactPackets and ControlPackets to update the state information for each client.

The visual setup is composed of a video wall with nine screens arranged in a 3 by 3 matrix and by local installation of client instances with adapted "Bouncing Ball" interfaces for desktop computers, touch screens and mobile PDAs. Each screen from the video wall is assigned to an instrument in the same order that they appear to the user in the PSOs Client interface. The clients are represented at the installation site as spheres with different colours, sizes and speed. Each client is assigned to a screen in the video wall which also limits the movement of the corresponding sphere, i.e., the limits of each screen are mapped to the limits of the PSOs Client's window. Whenever a new client connects, a colour is randomly chosen to represent their ball.

Two more parameters were chosen to provide visual feed-back: the speed of the ball and the events generated at the client's interface. Although there's an implicit visual feedback on the ball's speed, i.e., the sphere moves faster or slower on the screen, an additional feed-back was added by changing the saturation of the sphere's colour. Sometimes when the bouncing ball is set to a large size and occupies almost the whole screen it is hard to tell its speed because both a slow ball and a fast one will bounce a lot. Mapping the speed to the colour saturation – high saturation for a slow ball, low saturation for a fast one – helps viewers to distinguish these situations. When the Local Visual Representation Engine receives a packet, meaning that an

event was fired at the PSOs Client's interface, the client's sphere is temporarily turned into a polygon mesh representation.

The engine only has accurate information when clients send packets to the server. The rest of the time, the position of the bouncing ball has to be interpolated based on the information from the last packet. It is not possible to have a completely accurate representation of the user's bouncing ball due to network latency, different timing mechanisms on the clients and server, and because we cannot predict the user's actions. Despite all this, it is possible to get a fairly good representation of the various clients. The most noticeable representation artifact is the occasional “jump” of the sphere, e.g., sometimes the representation is changing more rapidly than the client's bouncing ball, so when a packet is received, the position is suddenly updated to the correct one causing the sphere to “jump” back.

## 5.4 Distinctive Software Features

The Public Sound Objects collaborative environment was designed to provide a graphical behavior-driven interface for individual sound manipulation. However the feed-back, regarding the performance of other present users in a given session, was meant to be exclusively auditory.

This topology is consistent with the Individual delayed Feed-back (IDF) introduced in Chapter 3, Section 3.2.3. The notion of whether a performance is more influenced by the auditory than by other forms of feed-back (visual or tactile) has been a topic of research and results are often bound to the specificities of the instrument (Dahl, S. and Bresin, R., 2001). Evaluation studies presented in the following section of this dissertation (5.5) led to the conclusion that most users found that, besides the auditory feed-back, it would be useful to have a graphic representation of other users in their individual GUI. Psychoacoustic mechanisms of musical grouping based on Gestalt and similarity rules of element's perception can play an important role in the musical Soundscapes of PSOs since each user is manipulating a distinctive Sound Object. Most research in Gestalt Theory has been concerned with the correlation amongst visual elements or amongst acoustic elements, and not so much with the relationship of visuals and sound (Temperley, D., 2001). However, from the Gestalt rules of proximity and similarity one can empirically infer that a user would make a strong association with moving objects that strike walls and corresponding triggered sounds, subconsciously associating each object with the respective sound and therefore gaining extra cues in the awareness of the other users' performance.



Based on these considerations, the latest version of the PSOs GUI released in May 2005 actually included the visual representation of all connected users.

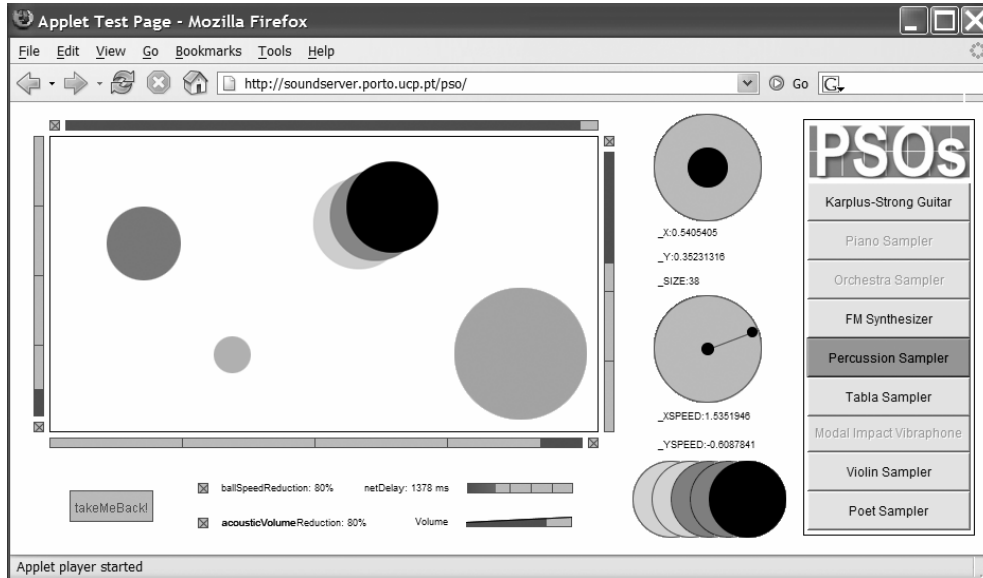


Figure 69. PSOs GUI version released in 2004; Including Multiple Users Graphic representation and other distinctive software features.

In this version of the PSOs GUI further enhancements were introduced.

1. The **selection buttons for each instrument** are now included in the same screen as the controller, allowing users to shift between instruments without leaving the main interface.
2. Checkboxes were appended to each wall fader. When checked, each a time a ball hits a wall, the pitch of the triggered sound is randomly assigned. This allows a totally algorithmic music generation environment.
3. A **volume fader** was added for the user's sound object on the global Soundscape. This fader is only accessible when the checkbox of **Acoustic Volume Reduction** is deactivated. The volume reduction is adaptive to real-time network latency measurements and it decreases with latency. The slower your connection, the lower is the volume of your Sound Object in the overall Soundscape (following a metaphor for sound volume and distance in real space)
4. A **Speed Reduction** Feature was added based on the notion of Latency Adaptive Tempo (LAT) introduced in Chapter 3, Section 3.2.1. The user can deactivate this

feature at the nearby Checkbox. With this speed reduction feature the bouncing ball will go as fast as the user's internet connection speed allows him.

The concept of Latency Adaptive Tempo (LAT) derived from an experiment and evaluation which demonstrated a direct dependency between Musical Tempo and tolerance to the disrupting effect of latency in a specific case of music collaboration (Standard Jazz performance). This concept was implemented in the PSOs providing a significant improvement on the awareness of individual performance in extreme conditions of latency on the auditory Feedback.

Even Though the PSOs interface has been designed as a simple musical controller with a High degree of abstraction in relation to a traditional musical instrument, it is important to realize the musical facets which are controlled by the user trough this interface are still the most basic musical parameters typically controllable in traditional music instruments: Pitch (the wall sliders), Tempo (The ball Speed and Size) and Dynamics (the volume slider).

In this sense this practical implementation reinforces the notion that LAT generally contributes to accommodate better a performer to his instrument when in the presence of Network delayed FeedBack.

In the case of PSOs changing Tempo is equivalent to changing a ball speed up to a certain extent. In this particular case a simple Ball speed reduction would not be enough since the disrupting effects of latency in the PSOs interface have particular side effects that need to be taken into account.

The following figure represents a temporal evolution a simple performance, in which an impact event occurs in  $t_n$  and the time lag between the impact and the reception on the client of the correspondent triggered sound generated at the central server is represented by  $\Delta t_n$ .

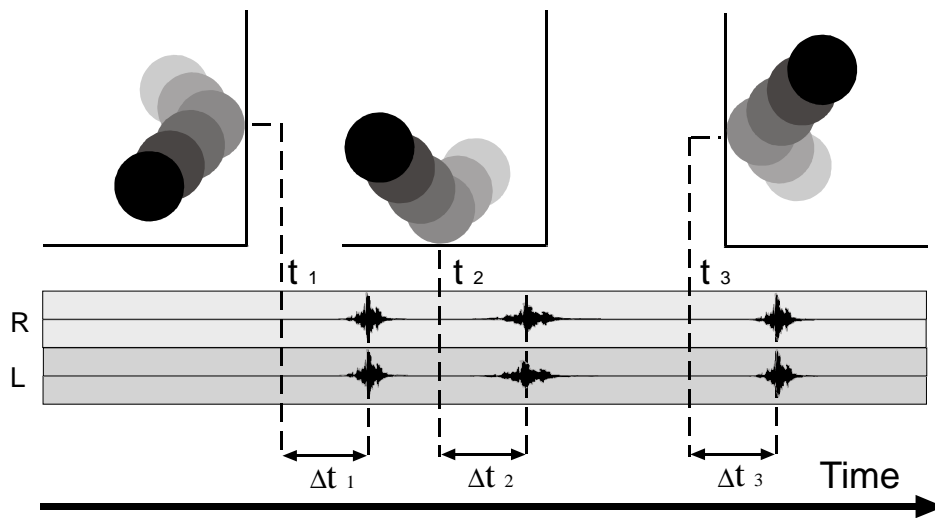


Figure 70. Representation of Impacts VS Triggered Sound without time lag overlap

From empirical observation of the system usage, it is clear that a user will perceived a sound played at  $\Delta t_n$  as being produced by  $t_n$ , as long as:

$$t_n + \Delta t_n < t_{n+1}$$

Thus, a triggered sound arrives to the client before another impact occurs, and the user can associate unmistakably each sound with the previous impact. However, if this condition is not met, a much more confusing scenario comes into play.

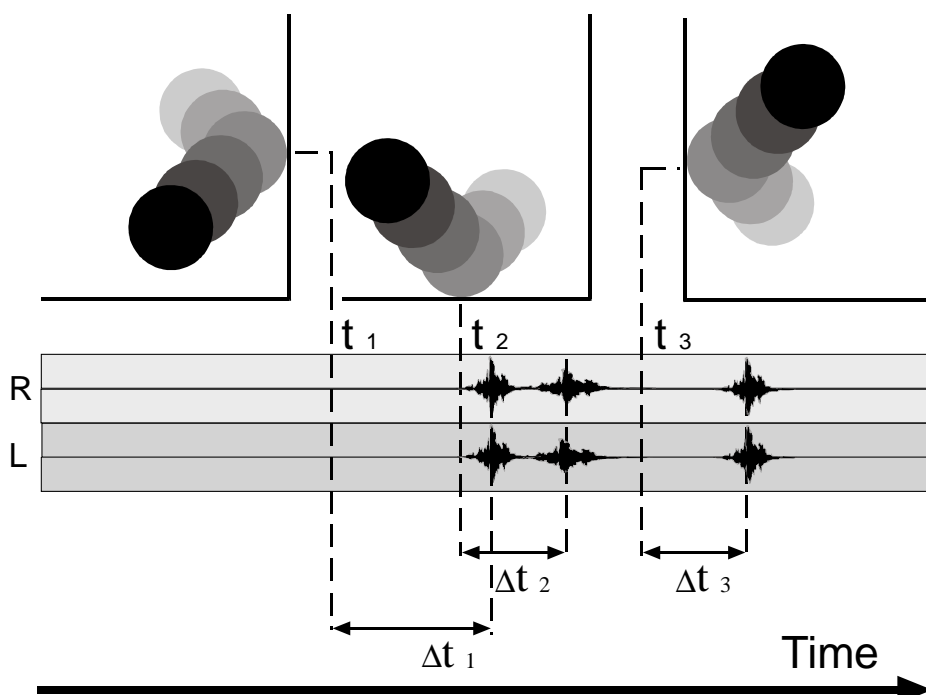


Figure 71. Representation of Impacts VS Triggered Sound with time lag overlap

The situation presented in the previous figure, creates a confusing perceptive correlation between an impact and a triggered sound, since a user can assume the first sound that occurs after the impact at  $t_2$  to be its corresponding Sound Object and the following sound at  $t_2 + \Delta t_2$  could be regarded as a sound from another user. An even worse scenario would be if  $t_n + \Delta t_n > \Delta t_{n+2}$ .

Therefore, in PSOs, the main criteria for a Latency Adaptive Tempo function is the condition  $t_n + \Delta t_n < \Delta t_{n+1}$ , instead of a limit defined by individual latency tolerance in instrumental performance, as described in Chapter 3, Section 3.2.1.

It should be clear though, that due to the ever-changing unpredictable nature of network latency, this adaptive process cannot always track these changes and adapt fast enough to attend to all impacts, however it minimizes the disrupting effect of time lag overlap.

An additional feature was implemented to improve perceptive correlation between an impact and a triggered sound, using a simple **sound panorama** adjustment at the sound server.

The basic idea consists of only transmitting a sound object through the Right Channel of the streamed Soundscape stereo mix, when a ball hits the right wall, transmitting only through the Left Channel when a ball hits the left wall and transmitting in both channels (L+R) if the ball hits the top or bottom wall.

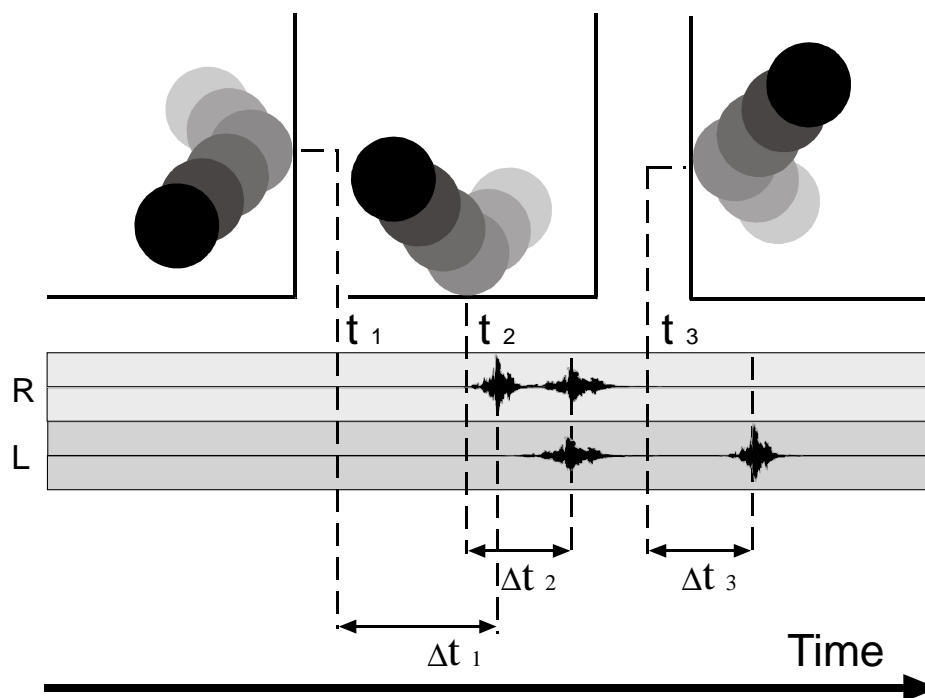


Figure 72. Representation of Impacts VS Triggered Sound with sound panorama adjustment

Sound Panorama Adjustment adds an extra cue to perception in temporal order of triggered Sound Objects and respective correlation to ball impacts.

## 5.5 System Evaluation

The Public Sound Objects System Evaluation has been conducted through different approaches. On one hand organized performances were carried out by different users, and specifically a performance taking place between Three Geographical displaced users was documented in video.

This performance took place on March the 31st of 2005, with users entering the System simultaneously from 3 distinct Locations: Toronto, Canada; Porto, Portugal; Barcelona Spain <sup>59</sup>.

---

<sup>59</sup> This performance is part of the PSO Short Video Essay presented as an Appendix to this dissertation and available on-line (MPEG4 QuickTime movie file) from the URL:

[http://artes.ucp.pt/docentes/abarbosa/pso\\_essay.mov](http://artes.ucp.pt/docentes/abarbosa/pso_essay.mov)

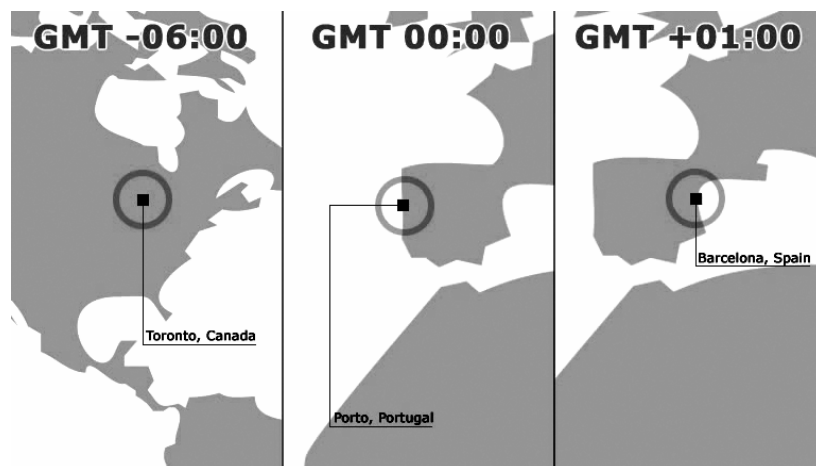


Figure 73. Locations of the PSO performance on March the 31st of 2005

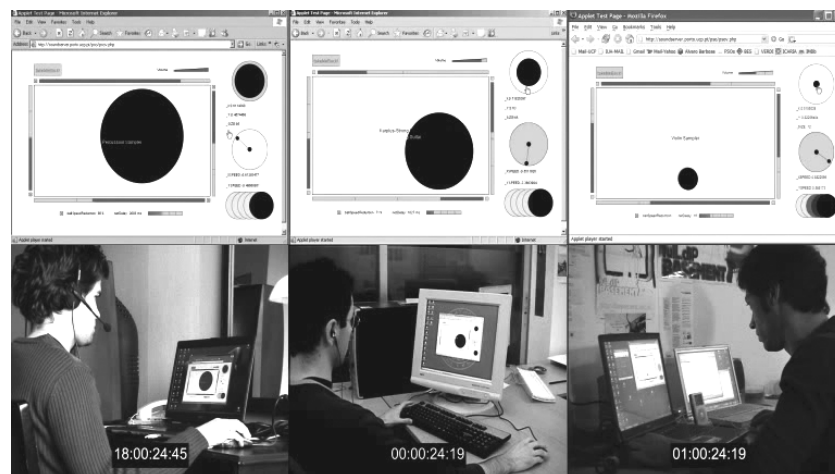


Figure 74. João Seabra, Jorge Cardoso and Álvaro Barbosa performing simultaneously with PSOs respectively in Toronto, Porto and Barcelona

The sound Server was located at the Porto Site and every user locally recorded their performance in video. Acoustic Communication was over a distance of: 5,648 kilometers (Toronto-Porto); 776 kilometers (Porto-Barcelona). From this experience a very strong feel of performative control was achieved by the users, even though communication latencies were in the order of hundreds of milliseconds.

Another source of empirical feed-back towards PSOs was the system presentation at two major international conferences on the Topic of Music Technology.



Figure 75. PSOs Installation at NIME 2005 – New Interfaces for Musical Expression Conference, 26-28 of May Vancouver, Canada.



Figure 76. PSOs Installation at ICMC 2005 – International Computer Music Conference, 5-9 of September Barcelona, Spain.

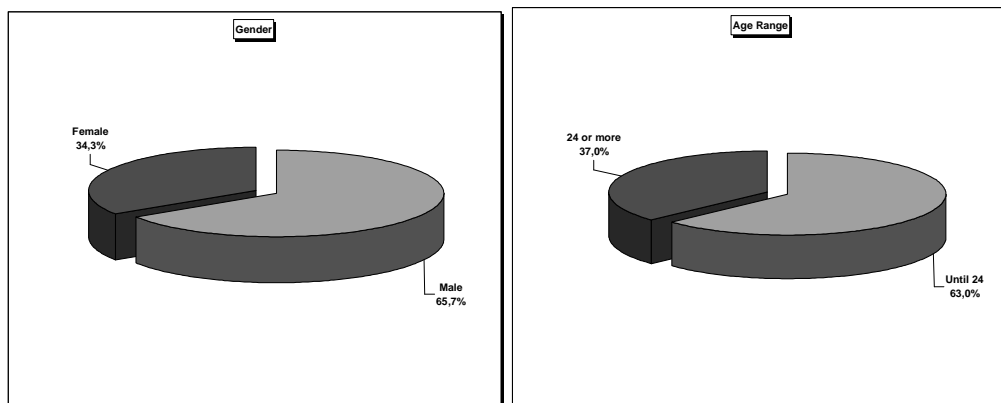
The main systematic evaluation process of PSOs was carried out while the complete system, including the physical setup at the server site, was installed at the Portuguese Catholic University Campus in Porto between 7 and 14 of October 2004.



Figure 77. PSOs Trial Installation at Porto School of Arts, 7-14 October 2004

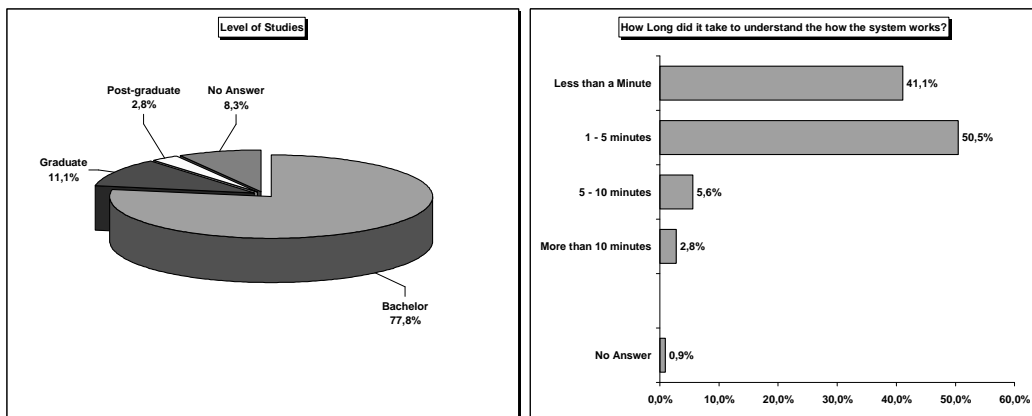
During this trial period several client instances were installed on campus and **109 subjects tested the system and answered questionnaires.**

The average results extracted from this opinion pool are presented in following Graphics:

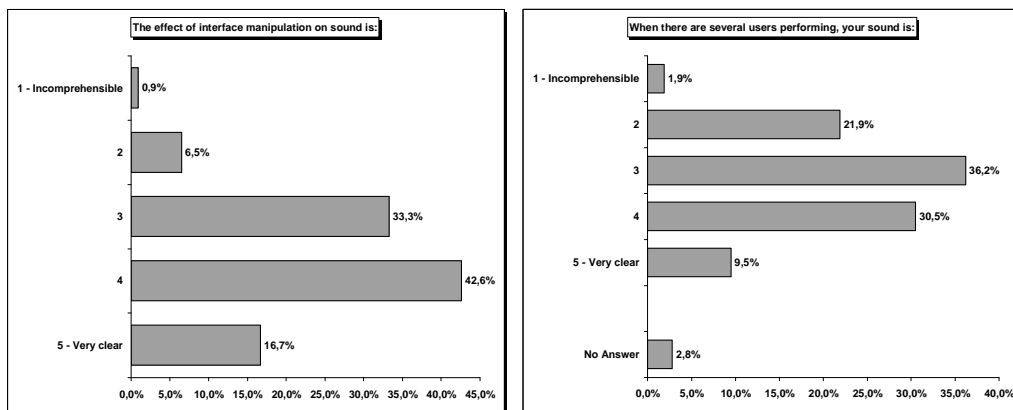


Graphic 4. Opinion Pool Characterization

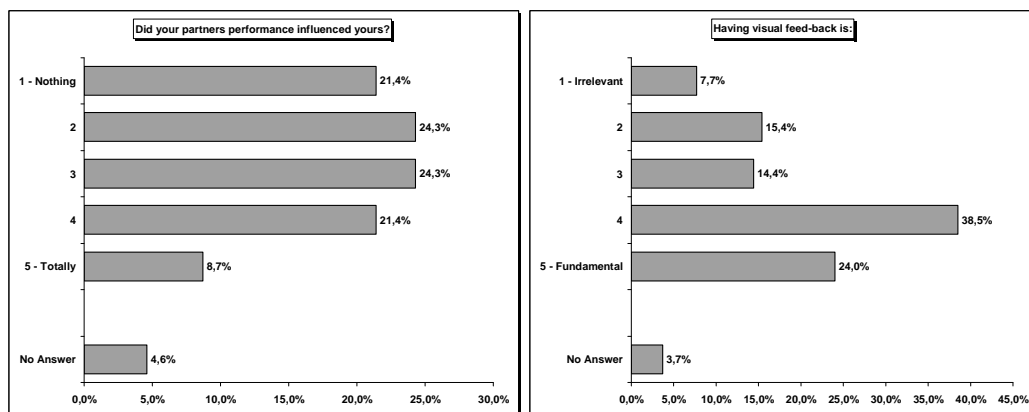




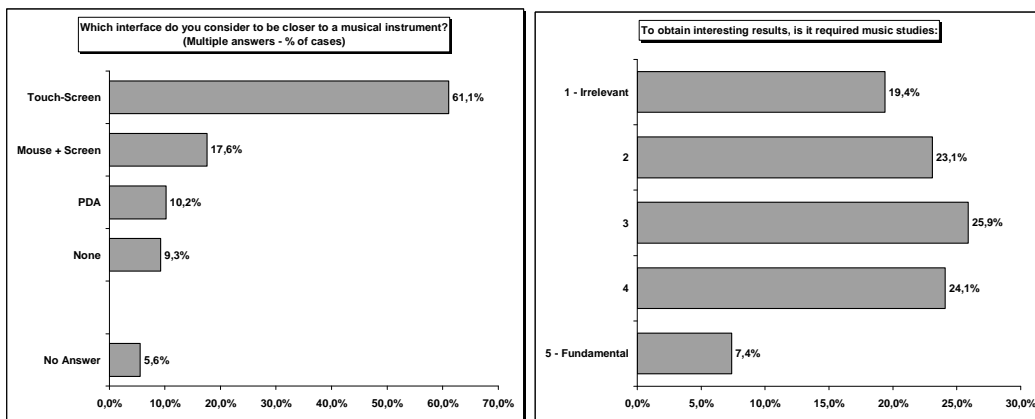
Graphic 5. Opinion Pool Characterization and question #1



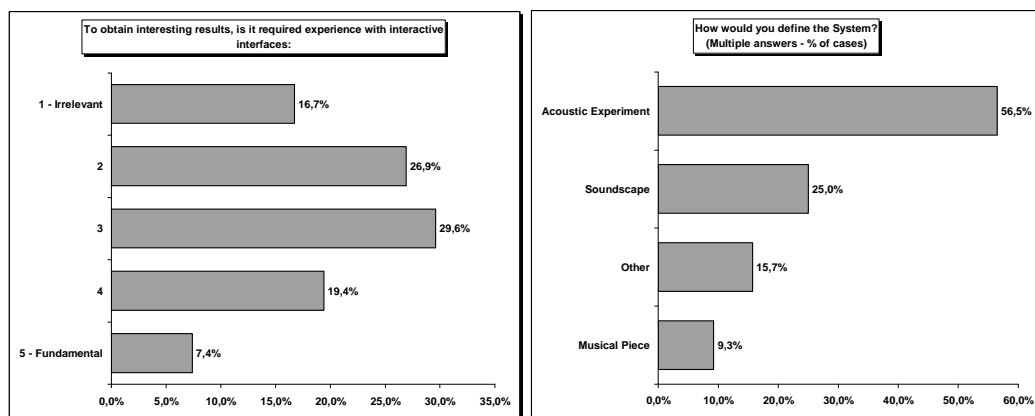
Graphic 6. Opinion Pool questions #2 and #3



Graphic 7. Opinion Pool questions #4 and #5



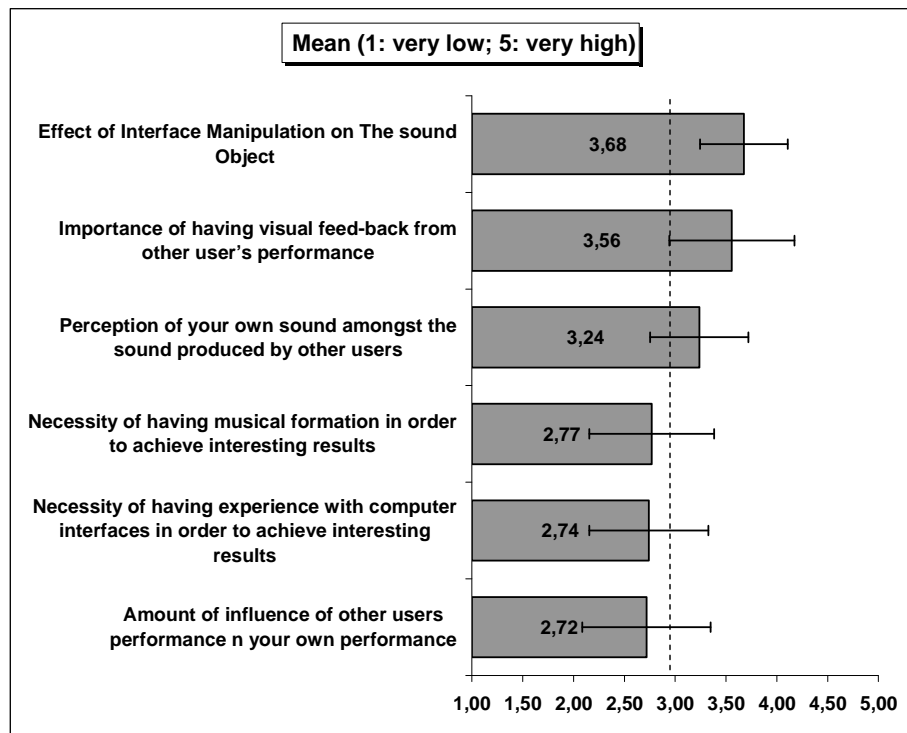
Graphic 8. Opinion Pool questions #6 and #7



Graphic 9. Opinion Pool questions #8 and #9

The following results met our expectations:

- (1) The interface is effective establishing a relation between the user action and its effect on the correspondent Sound Object;
- (2) The Sound Objects available at the current setup allow acoustic differentiation in the global Soundscape;
- (3) It is a system accessible to the general public, without requiring previous music education or previous GUI manipulation skills.



Graphic 10. Overall mean results from the opinion pool

When asked about further comments about the system most of the user's mentioned that the system is clearly about "experimental music" and the influence of other user's performances expressed by the following quotes: "The system only seems to make sense when used collectively"; "It is simple to be aware of the other users actions by looking into the video-wall and use it as reference when needed"; "Interacting with other users makes you achieve different results than you would get by yourself".

From these quotes and from the full statistical corpus it seems that in general users found the visual representation of other user's behavior useful, to enhance their own acoustic awareness.

## 5.6 Chapter Conclusion

For over two years the Public Sound Objects project has successfully functioned as an experimental framework to implement and test different approaches for on-line music communication.

The system is intended to be simple enough so that non-professional musicians can engage in a collaborative sonic performance, and in this sense the recent user study provided confirmation

that this goal has been achieved. More than half of the sample users considered that no musical training or experience manipulating computer interactive interfaces was required in order to achieve interesting results.

The System has a fast learning curve, since 90,6% of the sample users learned how to use the system in less than 5 minutes and 41,1 % of this group in less than a minute.

This was mainly due to the choice of experimental sound art as the musical aesthetic of the project and the fact that music is generated algorithmically, in the sense that even without a user's interference the "bouncing ball" will trigger an endless sequence of sounds with random tempo and Pitch.

In addition, even though Musical results of the system don't have familiar rhythmic or melodic structures, the control parameters which the "Bouncing Ball" interface present are Pitch, Tempo, Dynamics and Delay, which are very basic traditional music control parameters, and therefore will fit better into what a regular user would expect.

Further improvements are suggested by the evaluations of the sample users regarding visual representation of other users at the client instance, and in fact that was the impression one received by performing at the installation site, since it was inevitable to correlate the visual representation of other users and their acoustic contribution to the piece.

The Network Tempo Adaptive Latency and the coherent sound panning, implemented in the PSOs latest version, represent a significant improvement in the system usability regarding the disrupting effect of latency.

# Chapter 6

## Conclusions and Future Work

Digital Technology transformed the way Humans physically interact with Musical Instruments. It brought together Computer Science and Music into the field of Music Technology (Serra, X., 2005). This dissertation addresses a very specific aspect of this field, which concerns the usage of computer networks for collaborative music practices. The research carried out over the last five years departs from a basic understanding of the Computer Science field entitled Computer Supported-Cooperative Work (CSCW) and a broad analysis of the collaborative music practice domain (Jordà, S. and Barbosa, A., 2001). Throughout the development of this thesis the author narrowed down the focus of contributions, to this new and promising area, by progressively constraining his work to address very specific open questions presented on Chapter 1.

### 6.1 Summary of Contributions

#### **Survey and Classification**

Chapter 1 and 2 of this dissertation focus in getting Music and CSCW in perspective acknowledging the term **Networked Music** as any collaborative music paradigms approached in the context of computer mediated communication. By extensively surveying and classifying representative projects focused on musical practice over computer networks a major contribution is provided in terms of references and concepts for future work in this area.

It is possible to conclude from this study that the one of the most promising approaches to Networked Music is the possibility to create community oriented **Shared Sonic Environments**, where users can dynamically join and leave a group in a collaborative ongoing sonic performance based on the simple manipulation of sound, or even on the creation of musical structures. These systems go beyond the improvement of previously existing acoustic communication paradigms and focus on a breakthrough aspect of Internet collaboration, its

shared nature. These results led directly to the research and developments of the Shared Sonic Environment project **Public Sound Objects** introduced in Chapter 5.

### **Proposal of Models and Topologies**

With Sensor Technology and Computers a traditional musical instruments can be mimicked with interfaces designed to map human gestures to musical parameters in real time. However, the sound generation engine is now able to algorithmically create music, allowing an improvising performer to have a higher view on the awareness of musical development, by conducting the general direction of a musical behavior instead of producing music note by note as direct consequence of physical gestures. Digital Instruments not only allow behavior driven interaction, but also introduce complex interconnection possibilities between the instruments itself. Applied to a computer network, such instruments are accessible without geographical constraints. A model for an **Ubiquitous Virtual Music Instrument** is presented in Chapter 3, as well as high level models for wide-ranging Networked Music practices scenarios, which can combine different topologies based on centralized, distributed and peer-to-peer models, in order to meet diverse system' requirements.

More specifically it is presented a general model designed to support “Interface Decoupled Applications for Geographically Displaced Collaboration in Music” (Barbosa, A., Kaltenbrunner, M. and Geiger, G., 2003), the **Nomadic Music Instrument Model**, which served as a reference framework for Web Services and Acoustic Applications, as well as, the Public Sound Objects System presented respectively on Chapter 4 and 5 of this dissertation. In addition, is presented a brief introduction to **Multimodality applied to Networked Music**. The focus is on the Personal Digital Assistant (PDA) as client interface for Musical Applications (applied in the Public Sound Objects System) and the ReactTable Project as multi-user instrument developed with free software and openly accessible technology and materials.

### **Overcoming Constraints of Internet Acoustics**

When framed by a traditional music conception, performing music over the internet, requires real-time communication between performers. Yet, it is well known that communication latency (over a network or caused by intense computation) has a disrupting effect in musical synchronization. In such case musicians do not have an immediate feedback response to their physical gestures. The Acoustics of a performative space determines to a great extent the form and style of a musical performance. In chapter 4 is presented an analysis of perception issues related network Latency. The concepts of **Latency Adaptive Tempo and Dynamics** (LAT and

LAD) derived from an experiment and evaluation which demonstrated a direct dependency between musical tempo and tolerance to the disrupting effect of latency in a specific case of music collaboration (Standard Jazz performance). This concept has been implemented in the Public Sound Objects Project (PSOs) providing a significant improvement on the awareness of individual performance in extreme conditions of latency on the auditory Feedback. Even though the PSOs interface has been designed as a simple musical controller with a High degree of abstraction in relation to a traditional musical instrument, it is important to realize the musical facets which are controlled by the user through this interface are still the most basic musical parameters typically controllable in traditional music instruments: Pitch (the wall sliders), Tempo (The ball Speed and Size) and Dynamics (the volume slider). In this sense this practical implementation reinforces the notion that LAT generally contributes to accommodate better a performer to his instrument when in the presence of Network delayed Feedback.

Furthermore it was introduced the concept of and **Individual Delay Feed-Back** (IDF), based on the experimental demonstration that better latency tolerance is achieved if instead of having musician receiving direct acoustic feedback from their own instrument mixed with delayed feedback from the other performers, every musician listens to his acoustic feed-back delayed, together and in sync with the others. IDF has been applied to Alexander Carôt, Application for low-latency acoustic communication over the Internet entitled Soundjack (Carôt, A., 2004), and in the Public Sound Objects System.

### **Implementation of a Proof-of-Concept Application**

Most of the results previously mentioned converged in the implementation of the Public Sound Objects system (PSOs). The system is Shared Sonic Environment oriented towards the collective creation of musical structures. It is designed to be scalable and open to different implementations, both at the user interface level, parameters mapping and synthesis engine implementation.

In Chapter 5 the system is presented in full conceptual and technical detail, with special emphasis in distinctive software features:

- Latency Adaptive Tempo and Dynamics
- Individual Delayed Feed-back

- Behavioral Driven Interaction
- Panorama Cues for Temporal Order of Triggered Sound

A system evaluation of the PSOs system is also presented in Chapter 5, including different performance situations, as well as user's surveys on the system usage over different interface platforms (Desktop, Touch-Screen, PDA).

## 6.2 Future Directions

Some of the most immediate future work deriving from this dissertation is related with the Public Sound Objects System. It is by no means a finished project and new practical implementations will be conducted in the short term. The project will be undertaken by the Research Center for Science and Technology of the Arts (CITAR) from the Portuguese Catholic University and primarily the focus will be on the development of different synthesis engines and user interface layer to meet the ideas of Music Composers. Further specialization (panorama along the hit point) using wave field synthesis could also be performed at the installation side.

Nonetheless, some exploratory ideas and concepts resulting from 5 years of research in this specific field can be presented as a final note to this dissertation.

It is clear that technical aspects of Network Communication caused by Firewalls and security restrictions are increasing constraints for fast and effective acoustic communications, but hopefully there will always be methods to overcome it and meet the necessary quality of service for Music Applications. In the case of the Public Sound Object's project it was necessary to encapsulate all the communication data over the Http protocol in order to have a public access to the system, which resulted in a considerable increase of network latency. However, the introduction of Ross Bencina OSCGroups was a major step towards **Peer-to-peer real-time communication between musical systems**.

The notion of Latency Adaptive Tempo and Dynamics as well as Individual Delayed Feed-back, will soon be common concepts incorporated in the context of music performance on-line. Musicians will know that in order to jam on-line, they have to **improve their skills performing with delayed Feedback** and that they will be able to play as fast as their internet connection will allow them to.



Furthermore it wont be long until the development of **network audio drivers** will allow users to send and receive audio trough the network the same way Sound Card audio drivers allow this same procedure towards audio speakers, which are the form of sharing audio in the same physical space.

More and more will emerge musical applications that will **incorporate internet latency** as a functional part of the system instead of trying to eliminate it. This will possibly lead to music styles with **less rhythmic structures and** with **slow attack and decay sounds**.

Furthermore, the interfaces of such applications will be able to control sound generation engines that algorithmically create music, allowing the performer to have a more distant view in the development of musical structures. By **conducting the general direction of a musical behavior instead of producing music note by note as direct consequence of physical gestures** the disrupting effect of latency is highly reduced.

## Bibliography

- Amatriain, X. and Herrera, P. *Transmitting Audio Content as Sound Objects*. 15-6-2002. Proceedings of the AES22 International Conference on Virtual, Synthetic and Entertainment Audio.
- Amir, Y., Dolev, D., Kramer, S. and Malki, D. *Transis: A Communication Sub-System for High Availability*. 1992. Proceedings of the 22nd Annual International Symposium on Fault-Tolerant Computing (FTCS).
- Bafoutsou, G. and Mentzas, G. *A Comparative Analysis of Web-based Collaborative Systems*. 2001. 12th International Conference on Database and Expert Systems Applications.
- Bannon, J. *Discovering CSCW*. 1992. Proceedings 15th Information Systems Research in Scandinavia (IRIS) Seminar, Larkollen, Norway.
- Bannon, J. *CSCW - Challenging Perspectives on Work and Technology*. 1994. Proceedings of the Conference in "Information Technology & Organisational Change" Nijenrode Business School, The Netherlands.
- Barbosa, A., *Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation*. Leonardo Music Journal 13, 53-59 (2003).
- Barbosa, A., Cardoso, J. and Geiger, G. *Network Latency Adaptive Tempo in the Public Sound Objects System*. 2005. Proceedings the International Conference on New Interfaces for Musical Expression (NIME 2005); Vancouver, Canada.
- Barbosa, A. and Kaltenbrunner, M. *Public Sound Objects: A shared musical space on the web*. 2002. IEEE Computer Society Press. Proceedings of International Conference on Web Delivering of Music (WEDELMUSIC 2002) - Darmstadt, Germany.
- Barbosa, A., Kaltenbrunner, M. and Geiger, G. *Interface Decoupled Applications for Geographically Displaced Collaboration in Music*. 2003. Proceedings of the International Computer Music Conference (ICMC2003).
- Bargar, R., Church, S., Fukuda, A., Grunke, J., Kaislar, D., Moses, B., Novak, B., Pennycook, B., Settel, Z., Strawn, J., Wiser, P. and Woszczyk, W. *Technology Report TC-NAS 98/1: Networking audio and music using Internet2 and next-generation internet capabilities*. 1998. New York, Audio Engineering Society . AES White Paper; AESWP-1001.
- Bencina, R., Kaltenbrunner, M. and Jordà, S. *Improved Topological Fiducial Tracking in the reacTIVision System*. 2005. Proceedings of the IEEE International Workshop on Projector-Camera Systems (Procams 2005), San Diego (USA).
- Berners-Lee, T., *Weaving the Web*, p. 169, TEXERE London-New York, 2000.

## Bibliography

---

- Bischoff, J., Gold, R. and Horton, J., *Music for an Interactive Network of Microcomputers*. Computer Music Journal 2, 24-29 (1978).
- Black, U., *Internet Architecture an Introduction to IP Protocols*, Prentice Hall PTR, 2000.
- Blaine, T. and Forlines, C. *JAM-O-WORLD: Evolution of the Jam-O-Drum Multi-player Musical Controller into Jam-O-Whirl Gaming Interface*. 17-22. 2002. Conference on New Interfaces for Musical Expression (NIME).
- Bogazzi, R. P. and Dholakia, U. M., *Intentional social action in virtual communities*. Journal of Interactive Marketing 16, 2-21 (2002).
- Bongers, B., *An interview with Sensorband*. Computer Music Journal 22, 13-24 (1998).
- Brown, C. and Bischoff, J., *Computer Network Music Bands: A History of the League of Automatic Music Composers and the Hub*. In *At a Distance: Precursors to Art and Activism on the Internet*. (Ed. Annmarie Chandler and Norie Neumark) pp. 372-391, MIT Press, Cambridge, MA 2005.
- Bukofzer, M., *Music in the Baroque Era*, W.W. Norton & Co., New York 1947.
- Burk, P. *JSyn - A Real-time Synthesis API for Java*. 1998. Proceedings of the International Computer Music Conference (ICMC 1998).
- Burk, P. *Jammin' on the Web: A New Client/Server Architecture for Multi-User Musical Performance*. 2000a. Proceedings of the International Computer Music Conference (ICMC 2000).
- Burk, P. *Webdrum: 2000b*, SoftSynth. <http://www.transjam.com/webdrum/> , Accessed on June 21st of 2005
- Burk, P. *TransJam Server: 2005*, SoftSynth. <http://www.transjam.com/> , Accessed in June 21st 2005
- Cage, J., *Silence: Lectures and Writings*, pp. 57-60, Wesleyan University Press, 1961.
- Cano, P., Barbosa, A., Fabig, L., Goyon, F., Koppenberger, M. and Loscos, A. *Semiautomatic Ambiance Generation On-Line*. 2004. Proceedings of the 7th Int. Conference on Digital Audio Effects (DAFx'04), Naples, Italy.
- Cardoso, J., Carvalho, J., Teixeira, L. and Barbosa, A. *Soundserver: Data Sonification On-Demand for Computational Instances*. 2004. Proceedings of the Tenth Meeting of the International Conference on Auditory Display (ICAD 04), Sydney, Australia, July 6-9, 2004.
- Carôt, A. *Live Music on the Internet*. 2004. IT & design (FH Lübeck) - Germany.
- Chadabe, J., *Electric Sound - The past and promise of Electronic Music*, Prentice-Hall, Inc., 1997.
- Chafe, C., Gurevich, M., Grace L. and Sean T. *Effect of Time Delay on Ensemble Accuracy*. 2004. Proceedings of the International Symposium on Musical Acoustics, Nara - Japan (ISMA2004).

## Bibliography

---

- Chafe, C. and Leistikow, R. *Levels of Temporal Resolution in Sonification of Network Performance*. 2001. Proceedings of the International Conference on Auditory Display, Espoo - Finland (ICAD2001).
- Chafe, C. and Niemeyer, G. *Ping music installation, 2001*. 2001. Walker Art Center and San Francisco Museum of Modern Art - SFMOMA 010101. <http://www-ccrma.stanford.edu/~cc/sfmoma/topLevel.html>
- Chafe, C., Wilson, S., Leistikow, R., Chisholm, D. and Scavone, G. *Simplified Approach to High Quality Music and Sound over IP*. 159-163. 2000. Proceedings of the Digital Audio Effects Conference, Verona - Italy (DAFX2000).
- Chafe, C., Wilson, S. and Walling, D. *Physical Model Synthesis with Application to Internet Acoustics*. 2002. IEEE - Signal Processing Society. Proceedings of the International Conference on Acoustics, Speech and Signal Processing, Orlando - Florida (ICASSP2002).
- Chandler, A. and Neumark, N., *At a Distance: Precursors to Art and Activism on the Internet*, p. 343, MIT Press, Cambridge, MA 2005.
- Chowning, J., *The Synthesis of Complex Audio Spectra by Means of Frequency Modulation*. Journal of the Audio Engineering Society 21, (1973).
- Cooperstock, J. and Spackman, S. *The Recording Studio that Spanned a Continent*. 2001. IEEE Computer Society Press. International Conference on Web Delivering of Music (WEDELMUSIC 2001) - Florence, Italy.
- Cramer, F. *Combinatory Poetry and Literature in the Internet*. 2000. Forum Ästhetik digitaler Literatur, Universität GHK Kassel. Forum Ästhetik digitaler Literatur, Universität GHK Kassel.
- Csikszentmihalyi, M., *El fluir y la psicología del descubrimiento y la invención*, Paidós Iberica - Barcelona, 1998.
- Csikszentmihalyi, M., *Implications of a Systems Perspective for the Study of Creativity*. In *Handbook of Creativity*. pp. 313-335, Press Syndicate - Cambridge University, 1999.
- Curtis, P. *Mudding: Social Phenomena in Text-Based Virtual Realities*. 1992. Proceedings of the 1992 Conference on the Directions and Implications of Advanced Computing.
- Dahl, S. and Bresin, R. *Is the Player More Influenced by the Auditory than the Tactile Feedback from the Instrument?* 2001. Proceedings of the Conference on Digital Audio Effects (DAFX 01), Limerick - Ireland.
- Duckworth, W., *Making Music on the Web*. Leonardo Music Journal 9, 13-17 (1999).
- Duckworth, W., *Virtual Music: How the Web got wired for Sound*, Routledge, New York 2005.
- Feller, R., *FMOL Trio: Live at the Metronom*. Computer Music Journal 26, 110-112 (2002).
- Ferreira-Lopes, P., Dias, A. and Coimbra, D. *Music and Interaction: Consequences, Mutations and Metaphors of the Digital Music Instrument*. 2005. Proceedings of the 2º Workshop Luso-Galaico de Artes Digitais (ARTECH 2005), V.N.Cerveira - Portugal.

## Bibliography

---

- Fischinger, O. *Device for Producing Light Effects*. Application 182,669 [Patent 2,707,103]. 1955. United States Patent Office, Los Angeles Calif.
- Föllmer, G. *Crossfade - Sound Travels on the Web - Soft Music*. 2001. San Francisco Museum of Modern Art; ZKM - Center for Art and Media - Karlsruhe; Walker Art Center - Minneapolis ; Goethe Forum - Munich. <http://crossfade.walkerart.org/>
- Furukawa, K., Fujihata, M., and Muench, W. *Small Fish CD-ROM (Mac/PC)*: 1999, Hatje Cantz Verlag, Senefelder Strasse 12, D-73760 Ostfildern
- Gadol, S. and Clary, M. *Nomadic Tenets - A User's Perspective*. Sun Microsystems Laboratories, I. 1994. The SMLI Technical Report Series.
- Gang, D., Chockler, V., Anker, T., Kremer, A. and Winkler, T. *TransMIDI: A System for MIDI Sessions Over the Network Using Transis*. 1997. Proceedings of the International Computer Music Conference (ICMC 1997).
- Garnett, G., *The Aesthetics of Interactive Computer Music*. Computer Music Journal 25, 21-33 (2001).
- Geiger, G. *PDA: Real Time Signal Processing and Sound Generation on Handheld Devices*. 2003. Proceedings of the International Computer Music Conference (ICMC2003).
- Ghezzi, D., Gilbert, J., Smith, A. and Jacobson, S. *The Cassandra Project*. 1996. <http://www.nyu.edu/pages/ngc/ipg/cassandra/>
- Hajdu, G. *Quintet.net - A Quintet on the Internet*. 2003. Proceedings of the International Computer Music Conference (ICMC2003), Singapore.
- Hajdu, G. *Composition and improvisation on the Net*. 2004. Proceedings of the Sound and Music and Computing Conference (SMC 2004), Paris - France.
- Hargreaves, D., Miell, D. and MacDonald, R., *What are identities and why are they important?* In *Musical Identities*. Oxford University Press, Oxford 2005.
- Hernst, O., Gurle, D. and Petit, J.-P., *IP Telephony - Packet-Based Multimedia Communication Systems*, Addison-Wesley Pub Co, 1999.
- Hirsh, I., *Auditory Perception of temporal Order*. Journal of the Acoustical Society of America 31, 759-767 (1959).
- Jordà, S. *CD - La Fura Dels Baus - Sergi Jordà F@ust 3.0 FMOL*. 1998. La Fura Dels Baus.
- Jordà, S., *Faust Music On Line (FMOL): An approach to Real-time Collective Composition on the Internet*. Leonardo Music Journal 9, (1999).
- Jordà, S. *Digital Lutherie: Crafting musical computers for new musics' performance and improvisation (PhD Thesis)*. 2005a. Pompeu Fabra University - Music Technology Group.
- Jordà, S. *Multi-user Instruments: Models, Examples and Promises*. 2005b. University of British Columbia, Media and Graphics Interdisciplinary Center (MAGIC) 2005. Proceedings of the New Interfaces for Musical Expression (NIME-05), Vancouver, May 26-28, 2005.

## Bibliography

---

- Jordà, S. and Aguilar, T. *FMOL: a graphical and net oriented approach to interactive sonic composition and real-time synthesis for low cost computer systems*. 1998. Proceedings of COST G6 Conference on Digital Audio Effects 1998.
- Jordà, S. and Barbosa, A. *Computer Supported Cooperative Music: Overview of research work and projects at the Audiovisual Institute - UPF*. 2001. Proceedings of MOSART Workshop on Current Research Directions in Computer Music.
- Jordà, S., Kaltenbrunner, M., Geiger, G. and Bencina, R. *The ReacTable\**. 579-582. 2005. Proceedings of the International Computer Music Conference (ICMC 2005), Barcelona.
- Kahn, D., *Noise Water Meat: a history of sound in the arts*, MIT Press, Cambridge, Massachusetts 1999.
- Kaltenbrunner, M., Geiger, G. and Jordà, S. *Dynamic Patches for Live Musical Performance*. 19-22. 2004. Proceedings of the 2004 International Conference on New Interfaces for Musical Expression (NIME-04), Hamamatsu, Japan.
- Karplus, K. and Strong, A., *Digital synthesis of plucked string and drum timbres*. Computer Music Journal 7, 43-55 (1983).
- Konstantas, D. *Overview of a Telepresence Environment for Distributed Musical Rehearsals*. 1998. Proceedings ACM symposium on Applied Computing.
- Konstantas, D., Orlarey, Y., Gibbs, S. and Carbonel, O. *Distributed Musical Rehearsal*. 1997. Proceedings of the International Computer Music Conference.
- Koperski, K. and Han, J. *Discovery of spatial association rules in geographic information databases*. 47-66. 1995. Springer-Verlag. Proceedings 4th Int. Symp. Advances in Spatial Databases, SSD. Egenhofer, M. J. and Herring, J. R.
- kramer, G., Walker, B., Cook, P., Flowers, J., Miner, N. and Neuhoff, J. *NSF Sonification Report: Status of the field and research agenda*. 1999. National Science Foundation.
- Lago, N. and Kon, F. *The Quest for Low latency*. 33-36. 2004. Proceedings of the International Computer Music Conference (ICMC2004).
- Latta, C., *A New Musical Medium: NetJam*. Computer Music Journal 15, (1991).
- Loreiro, R. and Serra, X. *A web interface for a sound database and processing system*. 1997. Proceedings of the International Computer Music Conference.
- Marvin, L. E., *Spoof, Spam, Lurk and Lag: the Aesthetics of Text-based Virtual Realities*. Journal of Computer-Mediated Communication 1, (1995).
- Miller, G. A., *WordNet: A lexical database for english*. Communications of the ACM November 1995, 39-45 (1995).
- Moller, M., Henshall, W., Bran, T., and Becker, C. ResRocket Surfer: 1994, Rocket Networks. <http://www.resrocket.com/> , Accessed in April 17th of 1999
- Moore, F. R., *Elements of Computer Music*, Prentice-Hall, Inc., New Jersey 1990.

## Bibliography

---

- Mulder, A. *Virtual Musical Instruments: Accessing the Sound Synthesis Universe as a Performer*. 1994. Caxambu - Minas Gerais, Brazil. First Brazilian Symposium on Computers and Music.
- Nelson, T. EJamming: 2005. <http://www.ejamming.com/> , Accessed in June 27th of 2005
- Obraczka, K. *Multicast Transport Protocol: A Survey and Taxonomy*. IEEE Communications Magazine January [ ], 94-102. 1998.
- Obu, Y., Kato, T. and Yonekura, T. *M.A.S.: A Protocol for a Musical Session in a Sound Field Where Synchronization between Musical Notes is no guaranteed*. 2003. International Computer Music Association. Proceedings of the International Computer Music Conference (ICMC2003), Singapore.
- Oram, A., *Peer-to-Peer : Harnessing the Power of Disruptive Technologies*, O'Reilly & Associates, 2001.
- Perlman, M., *Unplayed Melodies: Javanese Gamelan and the Genesis of Music Theory*, University of California Press, 2004.
- Puckette, M. *Pure Data*. 269-272. 1996a. International Computer Music Association. Proceedings of the International Computer Music Conference, San Francisco (ICMC96).
- Puckette, M. *Pure Data: another integrated computer music environment*. 37-41. 1996b. Tachikawa, Japan. Second Intercollege Computer Music Concerts.
- Pulkka, A. *Spatial Culling of Interpersonal Communication within Large-Scale Multi-User Virtual Environments - Master of Science Dissertation*. 1995. Human Interface Technology Laboratory - University of Washington.
- Ramakrishnan, C., Freeman, J. and Varnik, K. *The Architecture of the Auracle: a Real-Time, Distributed, Collaborative Instrument*. 2004. Proceedings of the Conference on New Interfaces for Musical Expression (NIME 2004), Hamamatsu - Japan.
- Reese, G., *Music in the Renaissance*, W.W. Norton & Co., New York 1954.
- Roads, C., *Microsounds*, MIT Press, 2001.
- Rocheso, D. and Fontana, F., *The Sounding Object*, 2003.
- Rodden, T., *A Survey of CSCW Systems*. Interacting with computers - the interdisciplinary journal of human-computer interaction 3, 319-353 (1991).
- Schaeffer, P., *Traité des Objets Musicaux.*, Le Seuil, Paris, 1966.
- Schafer, M., *The Soundscape: Our Sonic Environment and the Turning of the World*, Destiny Books, Vermont 1977.
- Schuett, N. *The Effects of Latency on Ensemble Performance*. 2002. Stanford University.
- Serra, X. *A System for Sound Analysis/Transformation/Synthesis based on a Deterministic plus Stochastic Decomposition*. 1989. Stanford University.

## Bibliography

---

- Serra, X. *Towards a Roadmap for the Research in Music Technology*. 2005. International Computer Music Association. Proceedings of the International Computer Music Conference (ICMC 2005), Barcelona.
- Smuts, J. *Holism and Evolution*. 1926. Macmillan, London UK.
- Snow, J., *On the Mode of Communication of Cholera*, John Churchill, New Burlington Street, London: England, 1855.
- Spicer, M. *AALIVENET: An agent based distributed interactive composition environment*. 2004. Proceedings of the International Computer Music Conference (ICMC2004).
- Stelkens, J. *peerSynth: A P2P Multi-User Software with new techniques for integrating latency in real time collaboration*. 2003. Proceedings of the International Computer Music Conference (ICMC2003).
- Suchman, L. *Notes on Computer Support for Cooperative Work*. 1989. Dept. of Computer Science, University of Jyväskylä, SF-40100, Jyväskylä, Finland.
- Tanaka, A. MP3Q: 2000. <http://fals.ch/Dx/atau/mp3q/> , Accessed in May 12th of 2004
- Tanzi, D., *Observations about Music and Decentralized Environments*. Leonardo Music Journal 34, 431-436 (2001).
- Temperley, D., *The Cognition of Basic Musical Structures*, MIT Press, Cambridge MA, 2001.
- TONOS Company. TC8 - Music Collaboration Tool: 2001. <http://www.tonos.com> ,
- Truax, B., *Acoustic Communication*, Ablex publishing company, 1984.
- Trueman, D. *Reinventing the Violin*. 1999. Princeton University - Department of Music.
- Wang, G. and Cook, P. *ChucK: A Concurrent, On-the-fly Audio Programming Language*. 2003. Proceedings of the International Computer Music Conference (ICMC 2003), Singapore.
- Wang, G., Misra, A., Davidson, P. and Cook, P. *Co-Audicle: A Collaborative Audio Programming Space*. 331-334. 2005. Proceedings of the International Computer Music Conference (ICMC 2005), Barcelona - Spain.
- Weinberg, G. *The Aesthetics, History, and Future Challenges of Interconnected Music Networks*. 349-356. 2002. Proceedings of the International Computer Music Conference.
- Weinberg, G., *Interconnected Musical Networks: Toward a Theoretical Framework*. Computer Music Journal 29, 23-39 (2005).
- Weinberg, G., Aimi, R. and Jennings, K. *The Beatbug Network: A Rhythmic System for Interdependent Group Collaboration*. 2002. Proceedings of NIME 2002. Dublin: MLE.
- Weinberg, G. and Gan, S.-L., *The Squeezables: Toward an Expressive and Interdependent Multi-Player Musical Instrument*. Computer Music Journal 25, 37-45 (2001).



## Bibliography

---

- Weinberg, G., Lakner, T. and Jay, J. *The Musical Fireflies - Learning About Mathematical Patterns in Music Through Expression and Play*. 2000. Proceedings of XII Colloquium on Musical Informatics 2000. A'quila Italy.
- Winkler, T., *Composing Interactive Music*, pp. 21-28, MIT Press, 1998.
- Wöhrmann, R. and Ballet, G., *Design and Architecture of Distributed Sound Processing Systems for Web-Based Computer Music Applications*. Computer Music Journal 23, 73-84 (2002).
- Woszczyk, W., Cooperstock, J., Roston, J. and Martens, W., *Shake, Rattle and Roll: Getting Immersed in Multisensory, Interactive Music via Broadband Networks*. Journal of the Audio Engineering Society 53, 336-344 (2005).
- Wright, M. and Freed, A. *Open SoundControl: A New Protocol for Communicating with Sound Synthesizers*. 1997. Proceedings of the International Computer Music Conference.
- Xu, A., Woszczyk, W., Settel, Z., Pennycook, B., Rowe, R., Galenter, P., Bary, J., Geoff, M., Corey, J. and Cooperstock, J. *Real-Time Streaming of Multichannel Audio Data over Internet*. 47 [11]. 2000. Proceedings 108th Convention of the Audio Engineering Society.
- Young, J. *Using the Web for Live Interactive Music*. 2001. Proceedings of the International Computer Music Conference (ICMC 2001).
- Young, J. and Fujinaga, I. *Piano master classes via de Internet*. 1999. Proceedings of the International Computer Music Conference (ICMC 1999).

# Glossary

## Acronyms

**ACM** Association for Computing Machinery

**AES** Audio Engineering Society

**API** Application Programming Interface

**CCRMA** Center for Computer Research in Music and Acoustics, Stanford University in San Francisco, USA

**CIRMMT** Centre for Interdisciplinary Research in Music Media and Technology, McGill University in Montreal, Canada

**CSCW** Computer-Supported Cooperative Work

**DAFX** Digital Audio Effects

**DIVE** Distributed Interactive Virtual Environment

**EPT** Ensemble Performance Threshold

**FMOL** Faust Music On-Line Software

**IMN** Interconnected Music Network

**IP** Internet Protocol

**IRC** Internet Relay Chat

**ISDN** Integrated Services Digital Network

**ISMIR** International Conference on Music Information Retrieval

**LFO** Low Frequency Oscillators

**MIT** Massachusetts Institute of Technology

**MOO** Object-Oriented MUD

**MTG** Music Technology Group at Pompeu Fabra University

**MUD** Multiple-User Domain/Dungeon

**NIME** New Interfaces for Musical Expression Conference

**PD** Pure-Data

**ICMC** International Computer Music Conference

<b>PDA</b>	Personal Digital Assistant
<b>PSO</b>	Public Sound Objects
<b>SDI</b>	Serial Digital Interface
<b>SIG</b>	Special Interest Group
<b>SIGGROUP</b>	Special Interest Group on Groupware
<b>SVE</b>	Shared Virtual Environment
<b>TCP</b>	Transmission Control Protocol
<b>UCP</b>	Portuguese Catholic University, Porto , Portugal
<b>UDP</b>	User Datagram Protocol
<b>UPF</b>	University Pompeu Fabra, Barcelona, Spain
<b>VMI</b>	Virtual Musical Instrument
<b>VST</b>	Virtual Studio Technology created by Steinberg
<b>W3C</b>	World Wide Web Consortium
<b>WWW</b>	World Wide Web
<b>XML</b>	Extensible Markup Language
<b>SDM</b>	Sound Data Mining

## Index of Terms

<b>Broadcast</b>	- Pages: 7; 17; 34; 57; 85
<b>Co-Located Musical Networks</b>	- Pages: 42; 88
<b>Community Music</b>	- Pages: 3; 134
<b>Electroacoustic Music</b>	- Page: 23
<b>Individual Delayed Feed-Back</b>	- Pages: 49; 119; 153; 169
<b>Internet2</b>	- Pages: 26; 58; 63
<b>Latency Adaptive Tempo</b>	- Pages: 49; 118; 134; 154; 167
<b>Multicast</b>	- Pages: 62; 106
<b>Music Composition Support System</b>	- Pages: 41; 44; 51; 88

**Networked Music** - Pages: 14; 23; 40; 78; 98; 165

**Remote Music Performance Systems** - Pages: 41; 57; 88

**Shared Sonic Environments** - Pages: 41; 69; 71; 133; 149

**Shared Soundscape** - Pages: 18-22

**Shared Virtual Environment** - Pages: 11; 69; 76

**Sonic Art** - Pages: 3; 21; 32; 76; 133

**Sonification** - Pages: 26; 109; 127; 129

**Sound Object** - Pages: 71; 76; 132; 134; 142; 106

**Sounding Object** - Page: 141

**Soundscape** - Pages: 17-29; 66; 76; 119; 135; 143; 153; 162

**Unicast** - Page: 57

**Virtual Community** - Pages: 11-18

**Virtual Musical Instrument** Page: 41

# **APPENDIX A**

Published Work by the Author

## Papers in Peer-Reviewed Journals

**Barbosa, A. 2005.** “Public Sound Objects: A Shared Environment for Networked Music Practice on the Web” – Organised Sound, Volume 10 Issue 3 - Cambridge University Press, (in Print, December 2005). (OS: ISSN: 1355-7718)

**Abstract:** The Public Sound Objects (PSOs) project consists of the development of a networked musical system, which is an experimental framework to implement and test new concepts for on-line music communication. The PSOs project approaches the idea of collaborative musical performances over the Internet aiming to go beyond the concept of using computer networks as a channel to connect performing spaces. This is achieved by exploring the internet’s shared nature in order to provide a public musical space where anonymous users can meet and be found performing in collective Sonic Art pieces. The system itself is an interface-decoupled Musical Instrument, in which a remote user interface and a sound processing engine reside with different hosts in an extreme scenario where a user can access the synthesizer from any place in the world using the World Wide Web. Specific software features were implemented in order to reduce the disruptive effects of network latency, such as, dynamic adaptation of the musical tempo to communication latency measured in real-time and consistent sound panning with the object’s behaviour at the graphical user interface.

**Barbosa, A. 2003.** “Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation” – Leonardo Music Journal 13 - MIT Press, Cambridge MA (LMJ: ISSN 0961-1215; Vol.13: ISBN 0-26275392-8).

**Abstract:** The ubiquitous nature of communication in computer networks, firmly manifested in the Internet era, provided a context for the introduction of different collaborative tools widely accepted by the on-line community, such as textual chats, white boards, shared editors, video conference systems, shared spaces for the exchange of multimedia documents or even simple e-mail based collaborative systems. On the other hand, for the last decades artists have used cutting edge computer

technology to maximize the aesthetics and conceptual value of their work, either by enhancing the way they traditionally create, or by using technology as a medium itself for art expression. The idea of using computer networks as an element in collective artistic creation is no exception, since it provides particularly engaging possibilities to achieve stylistic and conceptual originality. Network Systems for Music and Sonic Art Creation emerged in the last few years, allowing geographically displaced creators to collaboratively generate shared SoundScapes. In this article the author presents a discussion about different system designs, ideas and concepts approaching this new interaction paradigm.

## Papers in Peer-Reviewed Conferences

**Barbosa, A.;** Cardoso, J.; Geiger, G. **2005**. “Network Latency Adaptive Tempo in the Public Sound Objects System” – Proceedings of 2005 International Conference on New Interfaces for Musical Expression (NIME 2005); Vancouver, Canada.

**Abstract:** In recent years Computer Network-Music has increasingly captured the attention of the Computer Music Community. With the advent of Internet communication, geographical displacement amongst the participants of a computer mediated music performance achieved world wide extension. However, when established over long distance networks, this form of musical communication has a fundamental problem: network latency (or net-delay) is an impediment for real-time collaboration. From a recent study, carried out by the authors, a relation between network latency tolerance and Music Tempo was established. This result emerged from an experiment, in which simulated network latency conditions were applied to the performance of different musicians playing jazz standard tunes. The Public Sound Objects (PSOs) project is web-based shared musical space, which has been an experimental framework to implement and test different approaches for on-line music communication. This paper describe features implemented in the latest version of the PSOs system, including the notion of a network-music instrument incorporating latency as a software function, by dynamically adapting its tempo to the communication delay measured in real-time.

Teixeira, L.; **Barbosa, A.;** Cardoso, J.; Carvalho, J.; and others **2005**. “Online data mining services for dynamic spatial databases II: air quality location based services and sonification” – Proceedings of the II International Conference on Geographic Information (GISPLANET 2005), Lisbon, Portugal.

**Abstract:** This paper introduces online data mining services for dynamic spatial databases associated with environmental monitoring networks. In particular, it describes an application that uses these services with sonification for air quality location based information services to the general public. The data mining services use Artificial Neural Networks, to



find temporal relations in the monitored parameters. The execution of the algorithms performed at the server side and a distributed processing scheme is used to overcome problems of scalability. In addition, two other families of web services are made available to support the discovery of temporal relations: vectorial and raster map? services and a sonification service. The map services were implemented in DM Plus, a client application presented in part I. The sonification service is described in this paper and illustrated through an application

Cano, P.; **Barbosa, A.**; Fabig, L.; Gouyon, F.; Koppenberger, M.; Loscos, A. **2004**. "SemiAutomatic Ambiance Genereation" – Proceedings of the 7th International Conference on Digital Audio Effects (DAFX 2004), Naples, Italy.

**Abstract:** Ambiances are background recordings of places used in audiovisual productions to make listeners feel they are in places like a pub or a farm. Accessing to commercially available atmosphere libraries is a convenient alternative to sending teams to record ambiances yet they limit the creation in different ways. First, they are already mixed, which reduces the flexibility to add, remove sounds or change the panning. Secondly the number of ambient libraries is limited. We propose a semi-automatic system for ambient generation. The system creates ambiances on demand given textual queries by fetching relevant sounds from a big sound effect database and delivering them into a sequencer multitrack project. Ambiances of diverse nature can be created easily. Controls are offered to the user to further specify its needs.

Cardoso, J.; Carvalho, J.; Teixeira, L.; **Barbosa, A.** **2004**. "SoundServer: Data Sonification On-Demand for Computational Instances" – Proceedings of the Tenth International Conference on Auditory Display (ICAD 2004), Sydney, Australia.

**Abstract:** The rapid accumulation of large collections of data has created the need for efficient and intelligent schemes for knowledge extraction and results analysis. The resulting information is typically visualized, but it may also be presented through audio techniques. Sonification techniques become especially interesting when the client application runs on

graphically limited devices such as mobile phones or PDAs (Personal Digital Assistants). In this paper we present an architecture for a sonification server that will be used in the Sound Data Mining project. In this project sound will be used to increase perception and present information extracted by spatial data mining techniques. The server is based on an audio synthesis engine and will relieve clients with little audio synthesis capabilities from the burden of sound processing. By providing sonification modules, this server can potentially be used on a variety of applications where sonification techniques are required.

**Barbosa, A.; Kaltenbrunner, M.; Geiger, G. 2003.** “Interface Decoupled Applications for Geographically Displaced Collaboration in Music” – Proceedings of the International Computer Music Conference (ICMC 2003), Singapore.

**Abstract:** In an interactive system designed to produce music, the sound synthesis engine and the user interface layer are fully integrated, but usually designed in parallel and in a modular way. Decoupling the interface layer from the synthesis engine, not only allows the use of best suited technologies and programming languages for each purpose, but also enhances the overall system flexibility. This paper discusses the idea behind a remote user interface and a processing engine that resides in a different host, taken to the most extreme situation in which a user can access the synthesizer from any place in the world using internet technology. This paradigm has promising applications in collaborative music creation systems for geographically displaced communities of user. The Public Sound Objects is an experimental system on which this concept is applied, and its currently under development at the Music Technology Group of the UPF in Barcelona.

**Barbosa, A.; Kaltenbrunner, M. 2002.** “Public Sound Objects: A shared musical space on the web” – Proceedings of International Conference on Web Delivering of Music (WEDELMUSIC 2002) - IEEE Computer Society Press, Darmstadt, Germany.

**Abstract:** In this paper we describe “The Public Sound Objects” project and its context. This project, which is currently under development, intends to approach the idea of collaborative musical performances over the Internet, going beyond most common paradigms where the network is

mainly used as a channel to provide a connection between performative spaces. At its final stage this system will provide a public performance space within the network where people can be found participating in an ongoing collaborative sonic event. The users connected to this installation are able to control a server side synthesis engine through a web-based interface. The resulting "Sound Objects" form a sonic piece that is then streamed back to each user. The user takes the role of a performer and his contribution has a direct and unique influence on the overall resulting soundscape. This ongoing event is also played back at the installation site in the presence of a live audience, with added contextual elements such as sound spacialization and metaphorical visual representation of the current participants.

Jordà, S.; **Barbosa, A. 2001.** "Computer Supported Cooperative Music: Overview of research work and projects at the Audiovisual Institute - U.P.F." - Proceedings of the Music Orchestration Systems in Algorithmic Research and Technology Workshop on Current Research Directions in Computer Music (MOSART 2001), Barcelona, Spain.

**Abstract:** In this paper the authors present an overview of recent work on ongoing projects at the Audiovisual Institute from the Pompeu Fabra University in Barcelona, focused on Internet collaborative virtual environments for music applications. Although presenting different strategies, they all put a special emphasis in performance and composition/production of music by groups of geographically dispersed communities of users, both in synchronous and asynchronous modes. It is presented an overview of the concepts and developments in: The FMOL project that approaches collaborative music composition and performance over the web in an asynchronous fashion in its original version (Jordà, 1998, 1999), and that is currently under development evolving to a new architecture with a synchronous paradigm and several relevant new features (Jordà and Wüst, 2001); The Public Sound Object project, which consists on a permanent web installation for collaborative musical performance, currently at preliminary development stage (Barbosa and Kaltenbruner, 2001).

## Other Related Publications

**Barbosa, A. (Editor) 2002.** “Musical Orchestration Systems in Algorithmic Research and Technology (MOSART)” – Chapter: Panel on Future Directions in Music Interfaces, Pages 302-306 – An EU IHP Network Project, General Editor: Kristoffer Jensen, University of Copenhagen 2002.

**Abstract:** In this paper is presented an overview and conclusions of topics and ideas discussed in the e Music Interfaces panel that took place during the MOSART Workshop – Workshop on Current Research Directions in Computer Music - Barcelona, November 17th of 2001. The invited members of Panel were: Antonio Camurri (DIST-University of Genova, Italy); Sergi Jordà (IUA-Pompeu Fabra University in Barcelona, Spain); Roger Dannenberg (Carnegie Mellon University, Pittsburgh, USA); Leonello Tarabella (CNUCE/C.N.R. in Pisa, Italy). The Chairman for the Panel was: Johan Sundberg (KTH-Royal Institute of Technology, Sweden). The panel had the duration of approximately one hour and it was structured in 3 parts: An introduction to the theme of the panel by the Chairmen; A five minutes introductory open statement by each one of the members; An open discussion on the introduced topics and ideas opened to the audience.