

**THE COMPOSER AS TECHNOLOGIST: AN
INVESTIGATION INTO COMPOSITIONAL PROCESS**

By

MILES SIMON WARRINGTON

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN in partial
fulfillment of the requirements for the degree

DOCTOR OF MUSIC (COMPOSITION)

Faculty of Humanities

UNIVERSITY OF CAPE TOWN

Date of submission: February, 2015

SUPERVISOR: MR. THEO HERBST

South African College of Music

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

DECLARATION

Miles Simon Warrington

I,, hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature:

Signed by candidate

 Signature Removed

18th May 2016

Date:

ABSTRACT

This work presents an investigation into compositional process. This is undertaken where a study of musical gesture, certain areas of cognitive musicology, computer vision technologies and object-orientated programming, provide the basis for a composer (author) to assume the role of a technologist and acquire knowledge and skills to that end.

In particular, it focuses on the application and development of a video gesture recognition heuristic to the compositional problems posed. The result is the creation of an interactive musical work with score for violin and electronics that supports the research findings. In addition, the investigative approach into developing technology to solve musical problems that explores practical composition and aesthetic challenges is detailed.

CONTENTS

ABSTRACT	ii
CONTENTS	iii
LIST OF FIGURES	viii
LIST OF TABLES	x
NOMENCLATURE	xi
ACKNOWLEDGEMENTS	xii
CHAPTER 1 - INTRODUCTION	1
1.1 - Overview of this Thesis	1
1.2 - Definitions	3
1.2.1 - Compositional Idea.....	5
1.2.2 - Compositional Problem.....	6
1.2.3 - Compositional Process.....	7
1.2.4 - Electroacoustic Music.....	9
1.2.5 - Gesture in Musical Context.....	11
1.2.6 - Hermeneutics.....	13
1.2.7 - Heuristics.....	13
1.2.8 - Psychoacoustics.....	14
1.2.9 - Technology.....	15
1.2.10 - Technologist.....	16
1.3 - Research Aim Outlined	17
1.4 - Research Questions Outlined	18
1.4.1 - Tools Needed.....	19
1.4.2 - Technological Requirements	20
1.4.2.1 - Non-Haptic Gesture Recognition System	20
1.4.2.2 - Electronic Music Generation System	20
1.4.3 - Compositional Result.....	21
1.4.3.1 - Compositional Elements	21
1.4.3.2 - Brief Outline of Composition <i>In-Gest</i>	22

1.5 - Research Rationale Outlined.....	23
1.6 - Contributions to the Field.....	25
CHAPTER 2 - THE COMPOSER AS TECHNOLOGIST.....	27
2.1 - Compositional Idea in Context.....	27
2.1.1 - Compositional Ideation and Technology.....	28
2.2 - HCI in Context.....	31
2.2.1 - Composition with Technology Specific to HCI: A Brief History with Examples.....	34
2.2.1.1 - HCI and Composer-Technologists.....	36
2.2.1.1.1 - Haptic Applications.....	38
2.2.1.1.2 - Non-Haptic Applications.....	40
2.2.1.2 - HCI and Composers Working with Technologists.....	41
2.2.1.2.1 - Haptic Applications.....	42
2.2.1.2.2 - Non-Haptic Applications.....	44
2.3 - Introduction: Technology as Heuristic and as the Composer's Tool	48
2.3.1 - Tools Identified.....	49
2.3.1.1 - Software.....	50
2.3.1.2 - Hardware.....	54
2.4 - The Next Step.....	56
CHAPTER 3 - GESTURE.....	57
3.1 - The Study of Musical Gesture.....	57
3.2 - Musical Gesture and Meaning.....	64
3.2.1 - Structural Outlines of Musical Gesture.....	64
3.2.1.1 - Gesture Space and Musical Gestures: A Contextual Structural Outline Pertaining to Compositional Process.....	65
3.2.1.1.1 - Gestural Units.....	69
3.2.1.2 - Functionality.....	72
3.2.2 - Summary.....	73
CHAPTER 4 - MUSIC AND MEDIATION TECHNOLOGIES.....	76
4.1 - Music Cognition and Psychoacoustics.....	76
4.1.1 - Music and Sonic Forms.....	78
4.1.1.1 - Mind and Matter.....	79
4.1.2 - Information Theory.....	83

4.2 - Mediation Technologies – Introduction	85
4.2.1 - Recent Research and Developments.....	86
4.2.1.1 - Selected Examples.....	88
4.2.2 - Towards a Mediation Technology Form for <i>In-Gest</i>	89
CHAPTER 5 - A MEDIATION HEURISTIC TAKES FORM	93
5.1 - Development of the Mediation Technology for <i>In-Gest</i>	93
5.1.1 - Devices and Capture Methods.....	94
5.1.2 - Gestural Information.....	94
5.1.3 - Processing.....	95
5.1.3.1 - Pure Data (PD).....	95
5.1.3.1.1 - Graphics Environment for Multimedia (GEM).....	98
5.1.3.2 - OpenCV.....	99
5.1.3.2.1 - Haarcascade Classifiers.....	100
5.1.3.2.1.1 - Training Method Outline.....	100
5.1.3.2.1.1.A - Collect Sample Data Sets.....	101
5.1.3.2.1.1.B - Determining ROI's of Positive Samples and vec File Creation.....	104
5.1.3.2.1.1.C - Train Using Haar-Like Features.....	109
5.1.4 - Summary.....	115
5.2 - Application of the Mediation Technology in <i>In-Gest</i>	118
5.2.1 - <i>In-Gest</i> PD Patches.....	119
5.2.1.1 - <i>In-Gest</i> Video Patch.....	119
5.2.1.1.1 - <i>In-Gest</i> Video Sub-Patches.....	121
5.2.1.2 - <i>In-Gest</i> Audio Patch.....	122
5.2.1.3 - Panic Patch.....	125
5.2.1.4 - Future Enhancements.....	126
CHAPTER 6 - THE COMPOSITION <i>IN-GEST</i>	129
6.1 - <i>In-Gest</i>: An Introduction	129
6.1.2 - Rationale Behind <i>In-Gest</i>	130
6.2 - <i>In-Gest</i>: The Composition	131
6.2.1 - Composing Phrases for Implied Musical Gestures.....	132
6.2.1.1 - Macro Structures: The Gestural Sheets.....	133
6.2.1.1.1 - Pitch Classes.....	133

6.2.1.2 - Micro Structures.....	134
6.2.2 - Generation of Electroacoustic Materials.....	135
6.2.3 - Scores in Process	138
6.2.3.1 - <i>In-Gest</i> : An Analysis.....	141
6.2.3.1.1 - <i>Walk</i>	143
6.2.3.1.2 - <i>Up</i>	147
6.2.3.1.3 - <i>Down</i>	151
6.2.3.1.4 - <i>Waltz</i>	155
6.2.3.1.5 - <i>Float</i>	159
6.2.3.1.6 - <i>Slow</i>	163
6.2.3.1.7 - <i>Fast</i>	165
6.2.3.2 - Determination of Tracking Data.....	168
6.2.4 - Synopsis	173
6.3 - Aesthetics: A Consideration.....	178
CHAPTER 7 - CONCLUSION.....	182
7.1 - In Summation of Processes, Procedures and Models.....	182
7.1.1 - Process Parts: A Summary.....	183
7.1.1.1 - Compositional Ideation.....	183
7.1.1.2 - Compositional Process.....	184
7.1.2 - Research Summary	187
7.1.3 - Composer–Technologist Transpired.....	191
7.1.3.1 - Compositional Process as Technologist: a Note from Personal Perspective.....	192
7.1.3.2 - Brief Summary of Observations from Tables 5., and 6.....	206
7.1.4 - Future Work	207
7.2 - Concluding Remark	208
APPENDICES	210
APPENDIX I - <i>In-Gest</i> Full Score.....	211
<i>Walk</i>	215
<i>Up</i>	216
<i>Down</i>	217
<i>Waltz</i>	218
<i>Float</i>	219

<i>Slow</i>	220
<i>Fast</i>	221
APPENDIX II - Audio Parent Patches	222
APPENDIX III - DVD ROM	225
DVD DISC CONTENTS AND FOLDER LAYOUT	225
Patches:	225
D:/IN_GEST	225
Demo Video Clips:	225
D:/DEMO_CLIPS	225
Audio Realisation:	225
D:/MIDI_VERSION	225
APPENDIX IV - Hand Model Contributors	226
REFERENCES	227

LIST OF FIGURES

Figure 1. – Compositional Process	8
Figure 2. – Gesture Study Schema	62
Figure 3. – Musical Gesture Signification Model A.....	66
Figure 4. – Musical Gesture Signification Model B.....	68
Figure 5. – Gesture Space and Gesture Units.....	71
Figure 6. – Complete Gesture Schema	74
Figure 7. – Music Mediation Model	85
Figure 8. – Mediation Technology Methodology.....	90
Figure 9. – <i>In-Gest</i> Mediation Model	91
Figure 10. – GEM Render-Chain	98
Figure 11. – Haar-Cascade Classifier Training Outline.....	101
Figure 12. – Positive and Negative Image Examples.....	104
Figure 13. – Training Sample Examples.....	108
Figure 14. – Stage in Final Trained Classifier.....	113
Figure 15. – Optimal Camera Setup	116
Figure 16. – Full Hand Finger-Tip Detection in PD/GEM.....	118
Figure 17. – <i>In-Gest</i> Main Video Patch.....	120
Figure 18. – Video Sub-Patch Network.....	122
Figure 19. – <i>In-Gest</i> Main Audio Patch.....	125
Figure 20. – Demo Recording Audio Setup and Layout.....	126
Figure 21. – Panic Patch Main Window.....	128
Figure 22. – Cascade Object Identifier Routes.....	128
Figure 23. – Pitches	134
Figure 24. – Thematic Linking in <i>Walk</i>	144
Figure 25. – <i>Walk</i> [detection 2P] Sub-Patches.....	146
Figure 26. – <i>Up</i> First Section Sonogram.....	148
Figure 27. – Section from [detection 1D] Sonogram Notation.....	148
Figure 28. – Notation of Shepard-Risset Glissando at [detection 3D] ...	150
Figure 29. – Phrase Variance in <i>Up</i>	151
Figure 30. – <i>Up</i> [detection 2P] Sub-Patches.....	152

Figure 31. – Rhythmic Linking Between [detection 2D & 4D]	154
Figure 32. – <i>Down</i> [detection 3P] Sub-Patches	154
Figure 33. – Electroacoustic Chord Structure in <i>Waltz</i>	156
Figure 34. – Rhythmic Variance in <i>Waltz</i>	157
Figure 35. – Motivic Interjections in <i>Waltz</i>	158
Figure 36. – <i>Waltz</i> [detection 7, 8, 11 & 13P] Sub-Patches	159
Figure 37. – <i>Float</i> [detection 1 & 2D] Main and Sub-Patch	160
Figure 38. – Wave-Form Time	161
Figure 39. – Thematic and Gestural Linking in <i>Slow</i>	163
Figure 40. – Augmentation of Appoggiatura Elements	164
Figure 41. – <i>Slow</i> [detection 3P] Sub-Patch	166
Figure 42. – <i>Slow</i> [detection 3P] Notational Elements	166
Figure 43. – Notation of ‘Doppler’ Effect	168
Figure 44. – <i>Fast</i> [detection 2, 3, 5, 6, 8, 9 & 10P] Sub-Patches	169
Figure 45. – GEM Detection Point Image Capture	171
Figure 46. – Data Range Control Sub-Patches	172
Figure 47. – Finger Data Correlation for <i>Fast</i> [detection 8D]	173
Figure 48. – Emergent Synthesis in <i>In-Gest</i>	175
Figure 49. – Directions of Movement in <i>In-Gest</i>	176
Figure 50. – Structural Scheme of <i>In-Gest</i>	177
Figure 51. – Compositional Process Summary	187
Figure 52. – Literature Review Topics and Proposed Models	190
Figure 53. – <i>Walk</i> Main Audio Patch	222
Figure 54. – <i>Up</i> Main Audio Patch	222
Figure 55. – <i>Down</i> Main Audio Patch	223
Figure 56. – <i>Waltz</i> Main Audio Patch	223
Figure 57. – <i>Float</i> Main Audio Patch	223
Figure 58. – <i>Slow</i> Main Audio Patch	224
Figure 59. – <i>Fast</i> Main Audio Patch	224

LIST OF TABLES

Table 1. – Gesture Mediation Technologies.....	87
Table 2. – PD Terminology	96
Table 3. – Training Attempt Historiography.....	114
Table 4. – Elements for Electroacoustic Material Generation.....	136
Table 5. – Electronic Works Prior to <i>In-Gest: A Dialectic Study</i> (1999 & 2000)	195
Table 6. – Electronic Works Prior to <i>In-Gest: gammaKhozi: Burning Ground & VLS Game</i> (2005)	196
Table 7. – Compositional Process: The Composer–Technologist of <i>In-Gest: Walk</i>	199
Table 8. – Compositional Process: The Composer–Technologist of <i>In-Gest: Up</i>	200
Table 9. – Compositional Process: The Composer–Technologist of <i>In-Gest: Down</i>	201
Table 10. – Compositional Process: The Composer–Technologist of <i>In-Gest: Waltz</i>	202
Table 11. – Compositional Process: The Composer–Technologist of <i>In-Gest: Float</i>	202
Table 12. – Compositional Process: The Composer–Technologist of <i>In-Gest: Slow</i>	203
Table 13. – Compositional Process: The Composer–Technologist of <i>In-Gest: Fast</i>	204

NOMENCLATURE

CV - Computer Vision

DVD - Digital Versatile Disc

GEM - Graphics Environment for Multimedia

HCI - Human Computer Interaction

HGR - Human Gesture Recognition

IEM - Institute of Electronic Music and Acoustics

IRCAM - Institut de Recherche et Coordination Acoustique/Musique

MAX - After Max Matthews

MIT - Massachusetts Institute of Technology

MSP - After Miller S. Puckette or Max Signal Processing

OpenCV - Open Computer Vision

PD - Pure Data

STEIM - Studio for Electro-instrumental Music

VGR - Video Gesture Recognition

ACKNOWLEDGEMENTS

The last three years of work would not have been possible without the enduring support of my mother and partner. Your patience, encouragement and tireless acceptance of suffering through the 'D.MUS-ness' has been without fault. I am eternally appreciative.

I wish to particularly acknowledge my friend Richard, without whom I would have been stranded on many an occasion. Your warm hospitality is greatly appreciated.

To my friends Toni, Shayne, Marc, Hardi and Billy who have provided accommodation when necessary - thank you all!

To my supervisor and mentor Theo Herbst, I am indebted to your insurmountable dedication and unfailing passion for our art. Without you at the helm, this would simply have not been possible. I am forever grateful for your trust and friendship. You are a doyen in our domain and have been a pillar of inspiration and vision for the future.

I wish to thank Meryl for your friendship and cheer. The time spent on our breaks outside gave me not only fresh air, but also a chance for a breather. I look forward to our future work!

I acknowledge Prof. Rebekka Sandmeier at the South African College of Music for providing me with a research space, without which I would not have been able to setup my studio and perform the task.

To Prof. Morné Bezuidenhout, for giving me the platform and space to present my initial ideas and research proposal.

I thank Dr. Martin Watt who has been a steadfast guide and a trusting and dependable source of knowledge and support.

To Larissa Johnson, I thank you very much for your patience and dedication in helping the realisation of this project. Without your violin work, this would not have been at all possible.

Without the critical funding provided by the following bodies, this degree would not have taken place:

- The National Research Foundation, Free-standing Doctoral Scholarship
- The Oppenheimer Memorial Trust
- The in-house doctoral scholarships at UCT, particularly the receipt of one of the doctoral Faculty of Humanities Research Associateships.

THANK YOU

CHAPTER 1 - INTRODUCTION

[...], the artist must also master fundamental disciplines such as certain branches of physics, mathematics, astrophysics, biology, and computer science. In reality, an artist is a theoretician, a manipulator, a creator of forms in movement. Seen from the point of view of art, all of our knowledge and our actions are but aesthetic expressions of forms and their transformations.

– Iannis Xenakis, Music and Architecture, 2008

1.1 - Overview of this Thesis

Throughout this thesis, the process of composition through a musical experiment that attempts to answer some compositional questions, themselves the fruits of ideation, is offered. The questions form part of a creative process, which contains problems that act as agents of inspiration at the start of a musical work, with the experiment resulting in said work that supports this thesis. In so doing, this solves the compositional problems.

The main purpose and aim is thus two-fold:

- document the relationship a composer (the author) can develop with technology in solving compositional problems which are concomitant to the process of composition and,
- detail the experiment that takes the composer on a journey of discovery through both the technological and musicological.

The above points are dealt with by performing a critical analysis of the process in developing and using technology suitable to answering compositional problems. This is achieved by embarking on a procedure of testing and knowledge acquisition to ascertain the best possible solution to the problems. This is done within a specific context, the context being where the composer can gain the most knowledge from the process where finding cost-effective ways to solve the problems are achieved. This method has two advantages:

- the process of the experiment affords the advantages of a composer acquiring knowledge, expertise and experience through a process of compositional problem-solving (where the composer *becomes a technologist* - definition follows), in contrary to the opposite practice where a composer *works with a technologist*¹;

¹ Whilst the former was deemed preferable over the latter and was the primary method adopted in this research undertaking, it was necessary to consult specialists in specific technological vocations for particular areas of the problem-solving, and these are detailed where necessary.

- the technology used, discovered and applied becomes a *heuristic* (definition follows), to the *compositional problem* (definition follows).

In order to provide an overview and context of the concepts and terms discussed, there follows a section of this chapter focusing on the definitions of these terms and concepts, to clarify their meaning and use in the context of this thesis.

1.2 - Definitions

This work finds itself rooted between the seminal debates focused on the problems of culture and technology found in Snow (1959) and Benjamin (1936). On the one hand it is argued that as artistic practice distances itself from its linkage to the "ritual" or "cult" status, the pure artifact touched by technology or other forms of reproduction, loses its "aura" and ability to communicate its "authority" and "authenticity" (Benjamin, 1936:II). On the other hand in sharp contrast, Snow (1959:17) states that where art and science merge "has been where some of the break-throughs came", and that "the clashing point of two subjects, two disciplines, two cultures—of two galaxies, so far as that goes—ought to produce creative chances." These differences in outlook are at the opposite ends of a debate in the art world where *artistic technique* is pitted against the *technical* or *technology*. In order to continue, an important distinction in

the use of the term *technique* is necessary here. *Technique* can refer to something that is learnt and adopted or applied or it can refer to a form of *technology*. For example, there are *techniques* of an artist such as brush strokes or the use of stretto by a composer. Then there are *techniques* that refer to technologies such as brush type or computer-assisted time-stretching respectively. The latter sense of *techniques* is adopted and serves as a descriptive term for the use of technologies in whatever applications are required (Becker & Eckel, 1994).

What seems to be misunderstood or perhaps underestimated, is the ability for the *technical* to enhance *artistic technique* to the point where the two benefit mutually, without the art losing its 'aura'. Indeed, the history of all art forms reveals important stages and periods that have adopted new "technical standard(s)" in an effort to evolve into "new art forms" (Benjamin, 1936:XIV). Yet, somehow if the dialogue between the *technical* and *artistic technique* is not each concomitant in the creativity taking place, the authenticity is lost. This is partly due to what Snow (1959:15) describes as the general view of literary intellectuals to "pretend that the traditional culture is the whole of 'culture', as though the natural order didn't exist." What is intimated with this is that there *is* no active communication, dialogue or understanding between the disciplines of art and technology or science. Yet, Martin Heidegger had in 1953, through the interpretation of the Greek term *techne* which derives its meaning from *Technikon (technology)*, proposed that art is itself a form of *technique*, a

view abutting the divides of the debate (Heidegger, 1977:13). The research here presents a position that adopts the ideals of both camps, and perhaps meets the middle ground proposed by Heidegger and also that of Snow's vision of mergence. This is undertaken to produce a work of art that embraces the *technical* and *technique*, with a view to achieving a form of syncretism within the 'two cultures' debate.

1.2.1 - Compositional Idea

For the purposes of this thesis, *compositional idea* refers to the thought or ontological process that starts with the identification of a *compositional problem* specific to the context of the thought at hand. The ontology referred to here, has its methodology based on that of phenomenology, principally that of Edmund Husserl (1859–1938), as detailed by Welch (1939), and is given detailed discussion in section 1 of Chapter 2. This methodology is also similarly applied where it is intimated that the "objects of one's mental activities" occur as "ideation, judgment, feeling and desire" (Siewert, 2011). The *compositional idea* is ultimately responsible for the ensuing *compositional process* and eventual resultant composition (see Figure 1., in section 1.2.3, for more details). Chapter 2 deals with *compositional ideation* in detail.

1.2.2 - Compositional Problem²

Essentially this is where compositional thinking and musical creation concerns itself with processes of cognition (ontological ideation) and the requirement for computational support or intervention. Tabor (1999:7) describes composition as a "task environment", and cites the theories of Newell & Simon (1972), who "define the task environment as an environment coupled with a goal, problem, or task—the one for which the motivation of the subject is assumed." *Compositional problem* in the context presented here is then certainly supported by (Truax, 1999:27), such that "sonology³ and musical thinking in general", are viewed "as a problem-solving activity". However, problem solving as such is not always a critical or intrinsic part of musical composition. This depends on the contents of the original *compositional idea*. If the idea contains problems that can be identified that require solving, then this would warrant the part(s) of the *compositional idea* being categorised as a *compositional problem(s)*. A good early example of this is demonstrated in composer Philippe Manoury's (1952) *Jupiter* (1987). The compositional problem was the requirement of score-following to allow for a chamber-music-like setting in electroacoustic music (McNutt, 2003:297), in which the performer and electronics form a type of communicative bond and interaction. To this end

² Note: This is not to be confused with the *aim* of this thesis, which is to investigate compositional process that casts the composer in the role of a technologist. This is made clear in section 1.3 of this chapter.

³ According to Laske (1975), *Sonology* is a "theory of sound structures that carry a meaning and hence are musical" (Truax, 1999:21).

a pitch-following system was created at the Institut de Recherche et Coordination Acoustique/Musique (IRCAM) using the MAX/MSP⁴ software program developed by Puckette (1988). Score-following can be achieved by running MAX on either a Macintosh or IRCAM Signal Processing Workstation (ISPW) together with the patch developed by Puckette for Manoury's piece (Settel, 2001:5, 44).

1.2.3 - Compositional Process

In spite of receiving its own section here, it should be noted that *compositional process* ought to be also recognised as a form of *compositional problem*. This is apparent since the *compositional idea* requires all of the sub-processes and their intrinsic problems, such as technological experimentation, engineering, the act of composition of musical ideas in the context of the problem itself, etc., in order to be realised.

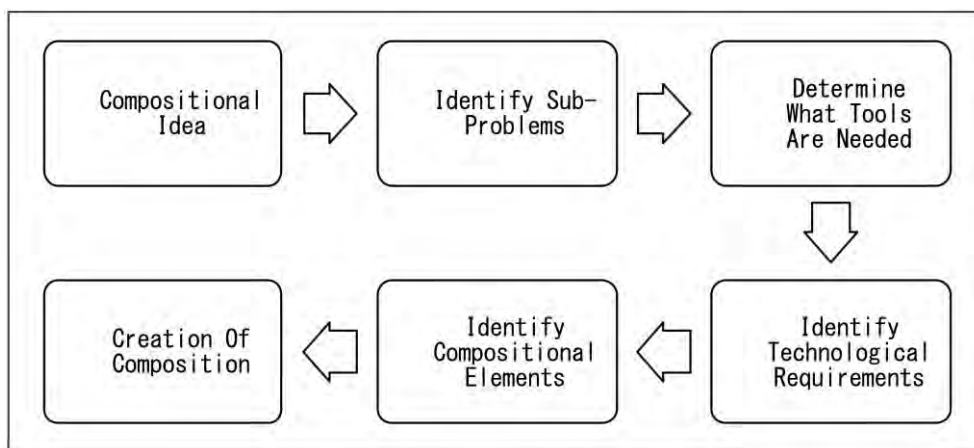
Throughout musical history, *compositional process* has been influenced by various techniques⁵. In context, *compositional process* was always characterised by the use of various techniques of the day, such as those seen as "new ways to treat musical material and become subject to historical development" (Becker & Eckel, 1994:8). Similarly, Metzger

⁴ For more information see <https://cycling74.com/products/max/> (accessed 2014, November 16).

⁵ As in section 2 of this chapter, *techniques* is a descriptive term for the use of technologies in whatever applications are required (Becker & Eckel, 1994).

(2009:4) describes a process of composition based on solution as looking to the past, where composers search previous pieces for “innovation” and “technical discoveries” and use these to solve the problems of previous works in their own. Certainly, the work created in partner to this thesis does look to previous examples for moments of inspiration. These processes form part of what Tabor (1999:3) describes as a “model-based” compositional process. This is in contrast to a “rule-based” compositional model that involves the composer navigating the creation of a work by “choosing his/her own processes to conform with the musical idea” (Tabor, 1999:4). The manner in which the composer (author) controls the way in which the problem-solving to the *compositional problem* is undertaken, is a form of ‘rule-based’ *compositional process* since the composition cannot be realised fully without this occurring. Figure 1., below, details the entire *compositional process* within the scope of this definition and demonstrates the ‘rule-based’ approach.

Figure 1. - Compositional Process



Thus the *compositional process* in the context of this thesis is adapted and influenced by both the techniques of the past (from a model-based) and from a 'rule-based' perspective where the design and implementation of a heuristic to the *compositional problem* informs the creative path.

1.2.4 - Electroacoustic Music

The period in the 20th Century 1948–1953 was paramount in planting the seeds for the development and evolution of electronic music to what we have in its present form (Cross, 1968:32). Within this period, there are two main forms. Demers (2010:27) describes Pierre Schaeffer's invention in 1948 of the technique that isolates "a specific sound by affixing it onto a disc that contained one locked groove (*sillon fermé*)". This appears to be the beginning of the techniques adopted to create a sound object (*objet sonore*). This signified the start of the *musique concrète* branch and led to the establishment of the Groupe de Recherches Musicales (GRM) in Paris.

In contrast to this, the 1951 Cologne electronic music (*elektronische Musik*) group made up of Stockhausen, Meyer-Eppler et al., formed (Demers, 2010:45). Differing from the French group, focus and attention was given to electronically generated tones to produce sine waves, white noise and other audio phenomena, which were then molded into compositions using the serialist principals of composers such as Schoenberg and Webern (Cross, 1968:33).

In New York during the same period, 'music for magnetic tape' hit the scene where composers such as Cage, Brown, Ussachevsky and Leuning utilised the principals and techniques of both the Paris and Cologne schools, whilst also incorporating transformations and processing of sounds recorded from acoustic instruments (Ibid).

Whilst Cross (1968) describes the grouping of all these forms as "electronic music", today they are more accurately collectively termed *electroacoustic music*. The principal reason for this, is that Hodgkinson (1986:1), details the differences between musique concrète and electronic music, eventually describing *electroacoustic music* as an entity "less concerned with the origin of the sound material than with what is done with it afterwards". Demers (2010:29) adds a further characteristic to the overall contextualisation of electroacoustic music by reminding us that the term *electroacoustic* suggests a combination of acoustic, purely synthesised electronic sounds and those in the Schaefferian model of manipulated recorded sounds. Thus, musique concrète and electronic music are divisions or subjects of *electroacoustic music*, since it relies on material generated or captured from these two parts. Demers (Ibid) refers to this as "post-Schaefferian electroacoustic music" and positions the genre between the purely musique concrète group and the aesthetic discourse of instrumental contemporary art music. Additionally, taking this into context, *electroacoustic music* is in this thesis, composed and

conceptualised using electronic means that is later “heard through loudspeakers” (Emmerson, 2007:xiii).

Moreover, due to the procedural audio manipulation capabilities of the system developed for the experiment, some of the music produced by electronic means is directly connected to and relies on parameters from the live performance, and should therefore be included in this definition. Thus in summation, *electroacoustic music* is the presentation of musical form that aesthetically aligns itself with western art music, and where transformed studio-generated and/or recorded material, and/or live interactive processing and manipulation are essential components.

1.2.5 - Gesture in Musical Context

Human gesture is a complex and vast subject because human beings have used movement to communicate since time immemorial. Gesture takes on the role of conveying information in the form of “feeling, idea, emotion, intention (and) expression” (Cadoz & Wanderley, 2000:29). This definition is however extremely general because it could apply to everyday tasks and is not necessarily specific to the musical domain. It is therefore necessary to apply focus to the term *gesture* for its use in the context of this thesis, particularly since the focus of the research deals with very specific gestures in the domain. There are only two broad kinds of *gesture* referenced in this work and their smaller sub-categories:

- *musical gesture*⁶: a compositional element or phrase that embodies some form of meaning or abstract communication through sound;
- *instrumental gesture*⁷: movement initiated by the performer through instrumental technique that results from the interpretation of the above *musical gesture*. These are further divided into smaller sub-categories, which are described in detail in Chapter 3.

Furthermore, Cadoz & Wanderley (2000:33), point out that there are three methods of gestural analysis:

- “A *phenomenological* approach (a descriptive analysis);
- A *functional* approach;
- An *intrinsic* approach (from musician’s point-of-view).”

Since an important part of this work is the attempt to explore the *hermeneutic* (definition follows) relationship between electronic-score and acoustic-score whilst simultaneously exploring the use of specific *instrumental gestures* to trigger *musical gesture*, the practice of applying all three approaches is undertaken. These are detailed in Chapters 4, 5 and 6.

⁶ After Jensenius et al., (2010:17,18).

⁷ After Cadoz (1988:5).

1.2.6 - Hermeneutics

In the context of this work, this term refers to the branch of philosophy that describes inter-textual relationships between concepts. Such that the description of a process that has such relationships could have multiple meanings. With the term applied here, it is perhaps more apt to use "*hermeneutic circle*⁸" where a "movement between the parts and the whole of the text" takes place (Ramberg & Gjesdal, 2014). The interactive nature of the composition itself and the processes as applied and detailed adhere to this definition. To illustrate this, a prime example of how the term *hermeneutics* is applied in this thesis, is much like the "metaphoric hermeneutics" as described in Filimowicz & Stockholm (2010:9), where the relationship between acoustic and electroacoustic score of affordances of *rising* for example, refer not only to pitch, but also to those of the increase in frequency of the waveforms present in the electronic score. Chapter 6 deals with *hermeneutics* as a process in the composition that partners this thesis.

1.2.7 - Heuristics

The use of the term *heuristic* applies in this thesis where a path to create a work of art is explored as the way in which the artist reaches the outcome (Livingston, 2013). Additionally, "A problem-solving

⁸ The term *hermeneutic circle* is explored in section 2.2 of Chapter 3.

heuristic is an informal, intuitive, speculative procedure that leads to a solution in some cases but not in others. Although there are several problem-solving heuristics, a small number tend to be used frequently. They are known as means-ends analysis, working forward, working backward, and generate-and-test" ("Heuristics", 2012, with emphasis). *Heuristic* development and context in this thesis is viewed thoroughly in Chapter 5.

1.2.8 - Psychoacoustics

Typically *psychoacoustics* is approached in this thesis from the perspective outlined by Loy (2006:150), where the subject is the mergence of acoustics and psychology, and similarly where the subject is further clarified as the "science of how we perceive sound" (Loy, 2006:154). Additionally where Chion (1983:16) wrote a guide to Pierre Schaeffer's seminal work *Traité des Objets Musicaux: Essai Interdisciplines* (1966), the description of *psychoacoustics* as the science that "looks for simple examples, such as pure frequencies, in order to study *the connections between variation in an elementary physical dimension of the object and variation in a sensory value*", is relevant. In essence, it is the study of how the physical properties of sound affect the senses and the resulting cognition and psychological interpretation of the audio data. Considering this definition, it is used in this thesis within the context of particular compositional elements conveying meaning (musical idea, thought or intention), which make use of specific techniques through the

application of mediation technologies. These are discussed in more detail in Chapter 4.

1.2.9 - Technology

Technology is defined by the Oxford English Dictionary Online as: “The branch of knowledge dealing with the mechanical arts and applied sciences; the study of this” (“technology, n.”, 2014). *Technology* in this context is used to describe the means of human environmental manipulation through the use and or development of tools specific to the *heuristic* of the *compositional problem*. This means the acoustic instrument (violin and its associated parts) may be ignored within the scope of this definition in the case of this study. In philosophy, *technology* is a phenomenon that is viewed from two broad perspectives. A “humanities philosophy of technology” that considers the “impact of technology on society and culture” and a philosophy of technology concerned with the understanding of technology itself where “both the practice of designing and creating artifacts [...] and the nature of the things so created”, is explored (Franssen, Lokhorst & van de Poel, 2013).

Technology as a form of *compositional process* is outlined in sections 4.1 and 4.2 of this chapter and later given detailed discussion in Chapters 2 and 5.

1.2.10 – Technologist

Technologist(s) is any person with a specialised understanding of a particular discipline of *technology*. The fields are wide-ranging and could include, electrical and/or mechanical engineering, computer sciences, robotics, software developers, etc. As the scope of the definition of *technology* in the previous section ignores the acoustic instrument, the historical significance of the influence of composers on the development of acoustic instruments is also subsequently ignored, since the *technology* dealt with here is electronically contextualised. Consequently this infers that the term *technologist* does not cover the development of the acoustic instrument itself.

Brün (1971) argues convincingly for the merits of composer-technologist collaborations in the development of new technologies. The premise is that through the act of composition, ideas formulated by composers can influence the creation of new technologies, which *technologists* should be critically aware of and that *technologists* themselves should “take a fresh view of the composer” (Brün, 1971:3). On the contrary, *composer-technologist(s)* are composers who have acquired similar skills to that of *technologists* within their specific technology field, usually to facilitate completion of musical projects requiring advanced technology. This thesis will focus on the latter context of *technologist* in the music compositional domain. The term *composer-technologist* is contextualised and discussed in detail in Chapter 2.

1.3 – Research Aim Outlined

Now that the definitions and their meanings in context with this research project have been detailed, the core research aim and objective can be described.

Whilst this has been briefly discussed in section 1.1 of this chapter, a broad summary is included here:

The principal aim of this research project is to investigate a compositional process where the investigator (composer) assumes the role of a technologist and/or collaborates with a technologist(s) in order to realise the compositional idea as defined.

Since this is extremely wide, the field was narrowed down, by focusing the investigation into an area of composition and electroacoustic music research that is emerging and where the composition of a work to test the research questions would contribute to the research field. Thus it became critical at the outset to ask very specific research questions that would act as knowledge and experience seeking agents in the process of solving the compositional problem as defined.

1.4 - Research Questions Outlined

From the above, the following questions were formulated in the very earliest stages of this project in order to ascertain what the compositional problems themselves would entail and thus start the overall compositional process:

1. Referring to the gestural theories of *sound-tracing*, *motor-mimesis* and *units* explored, (Jensenius et al., 2010), (Godøy & Leman, 2010), (Cadoz, 1988; Cadoz & Wanderley, 2000) etc., how can live instrumental gestures specific to a work (in this case for violin and electronics), inform the composition of musical phrases for such gestures?
2. Furthermore, with additional reference to the latter sources, how does the musical interpretation of the composed phrases inform the instrumental gestures themselves and, conversely, can the musical gestures themselves inform musical interpretation by the performer?
3. Is it possible for gesture recognition to enable electronic musical output either in the form of sound synthesis such as procedural audio or data-driven events akin to fixed audio content triggering, such as the examples explored by (Kimura & Risset, 2006, etc.; Overholt et al., 2009; Hayden & Kanno, 2011)?
4. Will this in turn create an 'interactive composition' using gesture recognition and interface as a heuristic?

It may be deduced from the above, that the overall problem can further be divided into the following areas, outlined (refer to Figure 1., in section 1.2.3 for path of the process), as follows:

1.4.1 – Tools Needed

- a. The determination of what kind of gesture recognition devices would be best in aiding the compositional realisation. Since the ideal realisation of the compositional idea will involve live musicians, a non-haptic⁹ or hands-free recognition system would be best. This would mean the use of a video recognition system.
- b. Determination of adequate computer ¹⁰ systems and output devices to enable realisation.
- c. Research into the most cost-effective means to develop the technological heuristic to the compositional problem.

⁹ 'Haptic' refers to the "Designating or involving technology (for entertainment, communication, etc.) that provides a user interface based on stimulation of the senses of touch and movement (kinaesthesia)" ("haptic, adj.", 2014). 'Non-haptic', refers to this process of interacting, but where there is no direct link between the senses of touch or movement, such as signal following or video tracking.

¹⁰ The use of 'computer' in this thesis refers to devices that fall within the following definition: "An electronic device (or system of devices) which is used to store, manipulate, and communicate information, perform complex calculations, or control or regulate other devices or machines, and is capable of receiving information (data) and of processing it in accordance with variable procedural instructions (programs or software); *esp.* a small, self-contained one for individual use in the home or workplace, used *esp.* for handling text, images, music, and video, accessing and using the Internet, communicating with other people (e.g. by means of email), and playing games" ("computer, n.", 2014).

These questions lead the research into areas involving in-depth testing and development of task-specific technologies to solve the problems. These are detailed in section 3.1 of Chapter 2.

1.4.2 - Technological Requirements

For the sake of further context in this chapter, it is necessary to give a brief outline of the technological requirements as discovered throughout the research process. They are listed below as:

1.4.2.1 - Non-Haptic Gesture Recognition System

- i. Capture devices - camera system;
- ii. Software - video recognition and tracking engine;
- iii. Interfaces - from video capture to computer software and from video recognition engine to (v.) below.

1.4.2.2 - Electronic Music Generation System

- iv. Input Data Recognition - dataflow from camera system to audio engine;
- v. Audio Synthesis Engines - Hardware/Software for sample playback and live processing;
- vi. Audio Graphic/Score Rendering System - allow musicians to read output in real-time and allow composer to assign specific musical tasks for interpretation of output.

These are further detailed in section 3.1 of Chapter 2.

1.4.3 – Compositional Result

Since the compositional problem-solving process has resulted in the creation of a composition, a brief outline of the composition itself is presented at this stage in order to contextualise the concluding section of the research rationale outlined below. However, a detailed chapter on the composition is presented in Chapter 6.

1.4.3.1 - Compositional Elements

The following compositional elements were identified in the initial stages of the overall process:

- a. Composition of approximately 15 minutes of original musical material for violin to allow for gesture recognition to take place. These musical ideas form the core structure of the work (see detail about the work *In-Gest* below), around which the interactive elements play a role in forming the overall composition.
- b. Creation of computer vision recognition system to track the finger-tips of the violinist.

- c. Analysis of overall process in realising the compositional idea where a critical analysis of compositional process was performed and documented.

1.4.3.2 - Brief Outline of Composition *In-Gest*

In the context of the previously defined terms *compositional idea*, *compositional problem*, technological *heuristic* and *gesture*, the work *In-Gest* - to 'take in' and 'gestural', is the result of the *composer-technologist* experiment. A contracted outline below serves to contextualise the work's details within the outline of the rationale that follows:

A computer and camera system was developed to act as a *heuristic* to the work and formed a critical part of the compositional process itself (section 4 of Chapter 5 deals with this in detail.) This tracks the finger-tips of the violinist and allows for the electroacoustic score to be realised in accompaniment to the acoustic score in a form of "invisible chamber music partner" (McNutt, 2003:297). It should be noted that an important result was discovered from this process, that being the phenomenon of a 'gesture interactive loop', where the performer and system are both experiencing information feedback, which determines performance and overall musical result.

1.5 - Research Rationale Outlined

One of the critical departure points for this research is supported by Herbert Brün's description of the principal of composers becoming technologists to aid the solution of compositional ideas in a heuristic sense: "The composer now defines technology as the science and art of applying knowledge to the desire for problem solving" (Brün, 1971:2).

This research undertaking is primarily concerned with compositional process that uses gesture heuristics. In order to place the rationale behind the project in context, it is necessary to discuss some historical context.

The Theremin from 1919 is arguably the forefather of musical gesture technology, since it was the first device to make use of a system to produce sounds without physical touch (Winkler, 1995:261). The Ondes Martenot invented in 1928 is a similar device to the Theremin. It is an electronic musical instrument controlled by the player wearing a sliding metal ring and keyboard. It produces alterable pitch of a single note by sweeping gestures of the right hand ("Ondes Martenot, n.", 2014). This is not yet 'gesture recognition' though, as this requires the movement to actually be interpreted, and this didn't happen until the first gestural recognition systems were developed in the 1960s (Myers, 1998:50). More specifically, video gesture recognition is a very recent and emerging

technology, and it is for this reason that this paper will focus on gestural technology that uses video human-gesture-recognition (HGR). Winkler (1995:264) also highlights the importance of composition informing "new media" and aiding the development of interactive technologies, and this was also used as a springboard in embarking on the research.

In more recent developments, Camurri et al., (2005:19) have explored various non-haptic and haptic gesture recognition techniques that involve musical performance to output specific technological tasks such as MIDI data, real-time score processing and sound synthesis, with a goal to demonstrating that "technology becomes a component of the artistic language."

Following on from the research undertaken by Camurri et al., (2005), there appears to be no distinct study of a composer-technologist process from both a technical and aesthetic view and where a cost-effective, non-haptic, video gesture heuristic was developed to explore the process and solve a set of compositional problems as proposed by the work *In-Gest* and this research project. In essence, the attempt to 'write music for the gestures themselves' is explored and undertaken within the context and analysis of the composer-technologist process.

In summation, the result is that the overall existing research looks at the broader subject of musical gesture in composed works (existing works

performed by the musicians and used as material for the experiments), but at the time this research was progressing, no study has as yet been undertaken to determine the effect of gesture recognition and interpretation on the process of composition and where this can act as a technological heuristic to realise a compositional idea. Furthermore, there appears to be a gap in the body of existing research where the process of writing music for the gestures themselves is used, and critically where the cast of the composer as technologist in the process of composition is explored.

1.6 - Contributions to the Field

There is much literature available on individual examples of how composers achieved their results using technologies old, new and yet undiscovered. As intimated previously in the *rationale outline*, there is not much written on the broader subject of the composer actually as the technologist or making use of said technologist(s) within the paradigm of video gestural recognition technologies used as a compositional heuristic. This is further supported by what seems to be a lack of research or examples where researchers have studied the compositional process of writing musical phrases for musical gestures themselves. And, critically, how through the creation of a 'gesture interactive loop' the composition of musical gesture is informed by instrumental gesture and vice versa.

The work *In-Gest*, which is partner to this thesis, has proven to be the 'compositional laboratory' testing the problem statement and rationale. Through its performance, testing and use of research material by other composers and researchers, it is hoped that it may be used to test their own compositional processes and explore 'the self' from the point-of-view of using technology as a heuristic to the overall process.

It is therefore intended that this study will help inform composers of the necessity to allow themselves an open mind and freedom of access to technology, using it as a pool of philosophy in the hope to broaden the possibilities of musical creation when applying technology as one of the indispensable tools or heuristic aids for compositional realisation. Furthermore, it should be noted that as far as non-haptic gesture recognition technologies and their application in musical composition is concerned, there appears to be no such other research being undertaken, presently on the African continent. It is for these reasons that the research is critical in providing an academic foothold in this particular area of music research in Africa.

CHAPTER 2 - THE COMPOSER AS TECHNOLOGIST

We must expect great innovations to transform the entire technique of the arts, thereby affecting artistic invention itself and perhaps even bringing about an amazing change in our very notion of art.

- Paul Valéry, Pièces sur L'Art, 1931 Le Conquete de l'ubiquite

2.1 - Compositional Idea in Context

Music is waiting to happen. The ontology¹¹ of the composer could be construed as a vesicle filled with the schema¹² of noise, ready to be organised. Similarly, Vaggione (2001:56) intimates that compositional ("musical") ideas exist in one's mind detached from what would soon be their organisation or assemblance. Put in a phenomenological perspective, the notion of the ontological process or 'phenomenon' behind the compositional idea is where the 'phenomenon' is "that which displays itself" (Welch, 1939:9). The "display" is that of the sound: the physical phenomenon of acoustics that can be heard and thus interpreted. Welch (1939:12) further breaks the notion of 'phenomena' into sections where

¹¹ Refer to section 2.1 of Chapter 1.

¹² As in Kant: "Any one of certain forms or rules of the 'productive imagination' through which the understanding is able to apply its 'categories' to the manifold of sense-perception in the process of realizing (sic) knowledge or experience" ("schema, n.", 2014).

"phenomena as essences are the contents of pure consciousness, and to this extent they become part of the consciousness to which they are presenting themselves."

But as with *technique* becoming a form of art and vice versa as discussed in section 2 of Chapter 1, the theory of ontological representations of *art* has in recent times¹³ been emancipated from the traditional approaches of the "metaphysical objective of expressing truth", and that "concentration has shifted to specific aspects of art, like the aesthetic use of signs, schematisation modes, or the message of art, and thus the metaphysical construct *art* has been separated into definable, explainable components" (Becker & Eckel, 1994:2). It is within these contexts of modern artistic theory that the notion of *compositional idea* is presented.

2.1.1 - Compositional Ideation and Technology

As described above, the idea consists of components. In the case of the compositional idea specific to this thesis, these components are formulated by asking questions, the questions being the fruits of ideation. In turn, these are concomitant to the research questions as outlined in section 1.4 in Chapter 1. But what *was* this idea? Originally, the core compositional idea *how to test the process of a composer becoming a technologist by writing a specific musical work to achieve that end*, was

¹³ Recent times should be considered in this context as the last 40 years, in parallel with the emergence of cognitive musicology.

presented. This idea is itself constructed of many different parts. They are outlined as:¹⁴

- the notion of technology incorporated in a musical work;
- the conception of musical phrases;
- this technology above being developed as part of the overall compositional process;
- the conception of an electroacoustic score to accompany the acoustic score and form the musical material that is dependent on the technology;
- all the above conceived as part of a compositional aesthetic within the electroacoustic art music paradigm.

The possibilities of the creation of a musical work from these questions are now made clearer. But the musical content is at this stage still abstracted. This is because the part of the compositional ideation, which specifically contains musical material, hasn't taken on any form until it is represented in sound. This is referred to as "representation by sound-image: a sonification of the conceptual idea", with the combining of concept and sound-image to form a "musical semiosis" (Nattiez, 1990:3).

Considering then that the compositional idea is a form of musical semiosis, this semiosis must also include the technology used or developed

¹⁴ Each of the above are described in more detail in the chapters that follow.

since it is part of the process involved in sonifying the idea. Thus it seems that technology, when part of a conceptual framework that results in musical semiosis, is productive or formative and therefore a form of poiesis, where "creative production, especially in the context of a work of art" ("poiesis, n.", 2014), presents itself. In contrast, part of the composition process is auditioning, where the composer may playback or replicate parts of the created material artificially. This may be viewed as a form of the partner to poiesis, namely that of esthesis, where "the perception of the external world by the senses", and thus also the reception of the musical semiotic takes place ("esthesis, n", 2014). This is a critical part of the compositional process once the compositional idea has been conceptualised and also serves to mold, adapt, morph or enhance the idea based on reference to the esthetic process. Thus it is fair to assume that when technology forms part of a compositional idea, it also forms part of the processes of creation and reception.

Formerly, art and technology have had their boundaries firmly encased within their own discourses, with their crossings even previously considered aesthetically misguided (Becker & Eckel, 1994:6). In this and the chapters that follow, technology is considered as an integral part of the creative process. This view attempts to demonstrate that the previously accepted boundary between art and technology is "brittle", and that the once "questionable concept of art as an autonomous social discourse, cannot be maintained" (Becker & Eckel, 1994:5).

In light of this, it is intended that technology as defined and considered so far, be understood as a critical component of compositional ideation, especially where the ideation itself relies on the formulation of said technology. Thus technology in this form of application, serves to cast the composer in the role of *composer-technologist*, where the technology as part of compositional ideation is tested, used and developed. To undertake an explanation of this concept, it is necessary to create a historical context of Human-Computer Interaction (HCI)¹⁵ in the musical domain and then detail some examples before exploring the process adopted to undertake the research questions in this investigation.

2.2 - HCI in Context

The global HCI domain covers many different facets of interactive technologies. In fact, it is a vast and interdisciplinary field. It not only encompasses object-tracking and recognition (the type of HCI this research is concerned with), but also vastly differing modalities such as those distinguished by Myers (1998). In the period 1963 to 1983, Myers identifies several important hardware and software tools mostly developed at universities with funding from government bodies, that had significant influence on technologies developed post that period. Categorised into groups described by single terms, they include "Drawing Programs, Text

¹⁵ For a definition of *computer*, refer to section 4.1 of Chapter 1. For a thorough definition and exploration of HCI, the reader is directed to <http://www.hcibib.org> (accessed 2014, November 20).

Editing, Spreadsheets, HyperText, Video Games, Gesture Recognition, Multi-media, Virtual Reality and Augmented Reality”, etc., (Myers, 1998:48–53). From these innovations, the interdisciplinary nature of HCI was able to grow, particularly until today where as far as the vision and object tracking sector is concerned, the “recognition and vision problem is highly interdisciplinary, spanning the fields of machine learning and decision making under uncertainty, robotics, signal processing, mathematics, statistics, psychology, neuroscience, HCI, databases, supercomputing and visualization/graphics (sic)” (Andreopoulos & Tsotsos, 2013:885).

Where a musical work involves computers that are directly interfaced with performers via some form of technological interface, then the system falls within the HCI sphere. The assumed and now standard model of HCI (a person interfacing with a computer via mouse, keyboard and visuals), are ignored. Musical applications of HCI are viewed as a highly specialised sector of HCI, “where the interaction between performer and a computing system engages several complex cognitive and motor skills” (Wanderley & Orio, 2002:66). Additionally, interactive computer music systems have the principle objective of creating “bi-directional communication between the performer and the computer” (Ibid). Owing to the complex nature of HCI in the musical domain, where multiple sensors and methods are adopted, it is viewed that this area falls into the Multi-Modal Human-Computer Interface (MMHCI) sphere. MMHCI is unlike the traditional mode of HCI (a one-to-one

interaction with a computer). In new purposes such as “remote collaboration, arts, etc.,” (Jaimes & Sebe, 2007:116), it allows interaction where they “are not always explicit commands, and often involve multiple users” (Ibid). Whilst MMHCI’s focus isn’t particularly that of music, it crosses it where its goal is to “determine how we can make computer technology more usable by people” (Ibid).

When applied in music, “multi-modality” refers to the targeted research area that is “concerned with the implementation of models and systems for detecting gesture(s) of the human body that becomes the natural interface able to give feeling and expressiveness to computer based multimedia performances” (Tarabella & Bertini, 2002:104–105).

Within the category of ‘gesture recognition’, specifically that of CV object tracking and recognition, the system proposed by Viola & Jones (2001) and adopted as discussed in section 2.2.2.1 of this chapter, places itself somewhere between the first vision-based system from 1995 that allowed “gestures to control the volume and channel functions of a television” (Karam, 2006:1), and the application proposed by Bhuyan, M.K., et al., (2012), which makes use of a Bayesian rule-based algorithm to locate finger-tips from colour segmentation. The ‘gesture recognition’ techniques adopted and developed in this investigation are the focus of a detailed discussion in Chapter 5, where they relate to the work created from the compositional process as outlined.

2.2.1 – Composition with Technology Specific to HCI: A Brief History with Examples

As intimated earlier, there exist at least two distinct methods for composers when working with technology as defined:

1. learn about existing technologies relevant to the task at hand or create and theorise the technology yourself i.e. become a technologist;
2. collaborate with a skilled technologist - an expert who is trained in engineering, computer science and programming, robotics, mechanics, audio-visual, cinematography and/or any other technological discipline that may be useful to the composer.

These two methods of process have been used for quite some time in musical history. The early stages of technological innovation with composition, was naturally devoid of HCI since computers had not yet been invented nor practical enough until the 1960s (as delineated by Myers (1988)) for any interactive use. The necessity to use or create new sounds and timbres e.g. Theremin and Ondes Martenot (as detailed in section 5 of Chapter 1), to the conceptualisation of new machines such as the *phonogène* to broaden the pallet of manipulated sound possibilities by Pierre Schaeffer and his technologist Jacques Poullin (Teruggi, 2007:217), are early examples of how the ideas of composers have developed or morphed technology. Later on, once the computer had appeared on the global scene, Pierre Boulez (1988),

founded IRCAM with the specific purpose of investigating the merge of electronics with composition and traditional instruments. Thus, should they need it, composers are often afforded the opportunity to work with technical and audio specialists (technologists). Notwithstanding, as time progressed, we now have examples of the real-time control of sound synthesis by non-haptic HGR, (to be detailed in Chapter 5), as well as bio-signal driven art. All of these demonstrate how music and technology have been, and continue to remain, evolving as partners in creativity.

Whilst the partnership between composers and technology has been meteoric, in many cases, the necessity for technology to help with the realisation of a composer's idea, may be a serious barrier to progress. This is because in some instances, the technology will simply be too daunting as it may not exist or is inherently extremely complicated. This is where the independent technologist becomes indispensable to the composer.

In light of this, the following examples detail some applications of HCI in the compositional domain within these two categories. These are in turn divided into haptic and non-haptic groups. In consideration of the definition of these categories as provided in section 4.1 of Chapter 1, haptic examples are viewed where the performer is in direct, visceral contact with a device that sends data to a computer system for musical interaction in a composition. Non-haptic refers to examples where the connection is electronically achieved by means of methods that do not

require direct sensing, such as video and audio signal communication with the computer system. It should be noted however, that there are examples in the electronic music repertoire that were conceived prior to the age of digital audio software. These emulate hardware setups that perform live processing of audio. A good example of this is Alvin Lucier's (1931) *I'm Sitting In A Room* (1970), where there now exist emulations in the form of Pure Data (PD)¹⁶ patches such as that by composer Johannes Kreidler (2013:204). Until the appearance of the PD patch, the example did not fall into the HCI domain, and thus was not conceived as a HCI composition. Similarly, devices where any input, be it audio signal, sensor information or controller data, which is not processed by a computer¹⁷ as defined, are ignored.

2.2.1.1 - HCI and Composer-Technologists

The examples that follow within this category, will demonstrate that where composers have assumed the role of technologists themselves, the focus of their creative work results in a process of technological innovation closely related to their artistic thinking. This is plainly evident from the scope and applications of the HCI interfaces developed, which are tailored specifically to the compositional ideas and problems at hand. The collaboration of artist and technologist, whatever the sphere, would usually occur between two or more people. However, in this case, it is

¹⁶ The reader is referred to Puckette (1997) for more details.

¹⁷ Refer to section 4.1 of Chapter 1.

internal and *personal*, taking on a form of ontological collaboration between differing fields of knowledge and experience occurring within the individual. One of the preeminent ways to critique or analyse this perspective, would be to undertake autoethnographic research of the process that results in the examples that follow. Autoethnography is a method of research that qualitatively analyses "personal experience in order to understand cultural experience" (Ellis, Adams & Bochner, 2011). In most of the cases found, no such material exists other than the output (work) of the process. However, the examination and critique of the compositional process that follows in Chapters 3 through 6 and particularly in Chapter 7 section 7.1.3, and to which this thesis finds its context and topic, fits into this model of research.

The use of technology in the compositional domain where developed by composers, has typically resulted in devices that lead to the extension of acoustic instruments to enhance their sonic possibilities. This is evident from the vast array of such applications in both the haptic and non-haptic groupings. These are referred to as *hyperinstruments*. Machover (1992), provides a detailed description of such extended systems, as tools that "transcend the traditional limits of amplifying human gestuality, and become stimulants and facilitators to the creative process itself." This is further contextualised where the "combination of machine-augmented instrumental technique, knowledge-based performance monitoring, and intelligent music structure generation, will lead to such creative

enhancement" (Machover, 1992:3). Thus the principal objective of augmented instruments is to provide scope for composers and performers through the use of the technique in composition, to amplify and extend their sonic capabilities. It is thus reasonable to deduce that when viewed within the description presented so far of technology's ability to be an essential component of both poietic and esthetic processes when applied to compositional ideation and procedure, that the *hyperinstrument* is an ideal vehicle for such representation. This therefore supports the choice in the original compositional idea as detailed, to create a work for augmented violin to demonstrate the approach. Consequently, the process of technological innovation by a composer, where the composer acts as or becomes a technologist, is supported.

2.2.1.1.1 – Haptic Applications

The early hyperinstrument work *Valis* (1987), by Tod Machover (1953), uses MIDI interfaces via computers to control the live synthesis of a score from "two hyperinstrument performers" at a piano and mallet percussion instruments (Rowe, 1993:75). Whereas *Begin Again Again* (1991) for 'Hypercello', by the same composer, uses an interactive system of sensors linked to a computer measuring physical movements of a cellist (Jessop, 2010:25).

Van Tonder (2004:24), used David Rokeby's (1960) *A Very Nervous System* (1995), (itself an early example of non-haptic interface - see section 2.1.1.2), as inspiration to model his own installation work called *Ephemeral Gumboots* (2004), which uses haptic-based touch sensors built-in to gumboots, worn and then used in a choreographed framework.

In recent developments, Donnarumma (2011) and Oritz (2012), have both explored the area of haptic systems that make use of bio-sensing instruments (mechanical myography - MMG), to measure physical properties of the human body in performance. These include, muscle tensions, heart-rate, temperature and even "biological body sounds", such as *Music for Flesh I* (2011), (Donnarumma, 2011:243). Miguel Oritz's *S&V* (2011), for saxophone and violin explores instrumental interface with a heart-rate monitor worn by the performers. The heart-rate monitor's data is sent in real-time to computers for processing of audio from the performers. Whilst this was developed with the hyperinstrument work of Machover et al., (1992) in mind, Oritz preferred the focus to "extend the performer" and not the instrument (Oritz, 2012).

In an interesting work that blurs the boundaries between the haptic and non-haptic groupings, Mari Kimura's (1962) work *Clone Barcarolle* (2009), uses a virtual gesture follower ("gf") controlled by accelerometers and gyroscopes. They are constructed into a glove the musician wears on the

right hand. This data is sent to a MAX/MSP¹⁸ patch running on a computer, which then “clones” the data with that of the audio signal from the musician (Kimura & Rasamimanana, 2012:n.p).

2.2.1.1.2 - Non-Haptic Applications

As pointed out earlier Van Tonder (2004), explores David Rokeby’s work *Very Nervous System* (1982–1990), in which he links the inter-actor to a video camera, an image processor, a computer, a synthesiser and a sound system. He succeeds in making a system in which the entire body can learn to move and “think” in the way that a pianist’s fingers “think” and respond more quickly than the conscious mind. In a *Very Nervous System* a computer observes physical gestures via a video camera and translates them into improvised music in real-time. The system “watches you”. The music is directly related to the qualities of the movements, creating a “direct and visceral relationship between body, space, sound and technology” (Van Tonder, 2004:22).

As a final example, Sam Hayden’s (1968) *schismatics* (2007, rev. 2010), explores live computer processing that makes use of “machine listening and learning techniques developed by Nick Collins” (Hayden & Kanno, 2011:486). The tools employed used both MAX/MSP and IRCAM’s OpenMusic¹⁹. One of the

¹⁸ Refer to section 2.2, of Chapter 1 for details.

¹⁹ See <http://repmus.ircam.fr/openmusic/home> for more information (accessed 2014, November 20).

substantial achievements of this work, particularly the 2010 version, is the more successful functioning of the computer “as an autonomous *virtual-improviser*, so there is less necessity for a second (human) performer to intervene to keep things musically interesting” (Hayden & Kanno, 2011:490).

2.2.1.2 - HCI and Composers Working with Technologists

In the mirrored process where composers work with technologists to achieve compositional realisation, the process is obviously less subjective or personal and requires communication with another party. Communicating artistic ideas can be a serious challenge to the artist, especially when the artist has no knowledge of systems or techniques that would enable the realisation of the creative idea. Instead, they must rely on methods of interaction across disciplines. Zhang & Cundy (2007), have studied the collaboration between artists and technologists using empirical data collection methods during artist-in-residency programs, to produce a digital artifact. Whilst one of the technologists in their data collection project made use of MAX/MSP, it wasn't for compositional purposes via the artist. However, they do list methods of interaction using a “computer-assisted communication mode”, which could be applicable across domains.

They are:

- the technologist(s) present what is done and describe how it was achieved;
- the technologist(s) and artist mutually tackle technical issues;
- the technologist(s) collaborate together to resolve the technical issues (Zhang & Candy, 2007:55).

Whilst these do not specifically cover the exemplar of composers working with technologists, the method and results of the qualitative analysis undertaken by Zhang & Candy (2007) could be useful to readers attempting collaborations with technologists.

2.2.1.2.1 - Haptic Applications

Composer and interdisciplinary artist David Rosenboom's (1947) *On Being Invisible* (1977), uses bio-sensing technology to capture electroencephalogram (EEG) data from a performer's head. This is applied to construct "complete musical forms" from the data, which are "analyzed (sic) by computer and used to direct the stochastic evolution of an adaptive, interactive electronic music system" (Rosenboom, 1997:25). The work was undertaken at Rosenboom's self-established Laboratory of Experimental Aesthetics at York University in Toronto in 1970, where projects over a seven-year period culminated in the realisation of *On Being Invisible* (Rosenboom, 1997:11).

1983 saw the invention by composer Michel Waisvisz (1949–2008) of the *Hands* device (Bongers, 2000:126). Co-developed with a team of researchers and engineers at the composer's STEIM²⁰ Studios in Amsterdam, the system consists of a cylinder for each hand to which several sensors and switches are attached in "rows of four keys; these provide pitch control (MIDI key-on, key-off) within one octave" (Waisvisz, 1985:314). The devices themselves are linked via cables to a personal computer running the "Control Signal Processor" (Waisvisz, 1985:315). Providing visual reference to connections with electronic systems in electronic music performances, lends, according to Emmerson (2007:137), "theatrical dimension to a performance."

Composer-performer Laetitia Sonami (1957), was the first artist to receive Bert Bongers's *Lady Glove* (1994) and subsequently a later model in 2001 (Bongers, 2000:152). These examples had the sensors sewn onto the upper-side of a "thin black mesh, arm-length, lycra glove tailored in Paris" (Sonami, 2010). The sensors themselves are a special kind of 'bend sensor' that is ideal for "tracking the rotational movement of the joints of the fingers" (Bongers, 2000:151). In 1999, composer-performer Walter Fabeck (1972) had a similar glove designed and built whilst at the Institute of Sonology²¹. Both these systems were sponsored by STEIM and run to the

²⁰ See <http://steim.org> for details (accessed 2014, November 20).

²¹ Refer to <http://www.sonology.org> for more information (accessed, 2014 November 23).

Sensorlab²² system, which maps the signals into a computer running a MAX/MSP patch.

2.2.1.2.2 - Non-Haptic Applications

In the early days of computer-assisted music, Pierre Boulez's (1925–2016) *Repons* (1981) for 24 musicians and 6 soloists uses the 4X system to transform and route soloists to loudspeaker systems (Boulez & Gerzso, 1988:47). Boulez made use of technologist experts at IRCAM using the 4 X computer system they developed, which used revolutionary real-time digital signaling processing for that time.

Detailed by May (2006:147), Philippe Manoury's (1952) seminal work *Jupiter* (1987), applies score-following by means of signal analysis and processing, which is then used to create a dynamic interactive performance between flute and computer. Working with the development team of the ISPW and realised at IRCAM, the system has since been ported to both MAX/MSP and a PD²³ patches running on any Apple Macintosh computer, which came with later versions up to 1996 and then again as recently as 2008.

²² Refer to <http://steim.org/product/discontinued-products/> for more information (accessed, 2014 November 23).

²³ For the latest up-to-date version see <http://msp.ucsd.edu/pdrp/latest/files/manoury-jupiter/> (accessed, 2016 April 30).

In a similar vein, Jean-Claude Risset's (1938) works *Variants* (1995) and *Violoné* (2006), employ signal processing and score-following using MAX/MSP to extend the violin (Kimura & Risset, 2006:407,408).

The work of composer-performer Jonathan Impett (1956) on his *meta-trumpet* (1994) project, highlights the role technologists can play in developing composition and performance systems for musical extension. Here the software "processes the incoming data and provides a composition and scheduling environment in which the composer can determine the system's response" (Impett, 1994:147). The aim of the *meta-trumpet* project describes a process where the instrument exists in an intimate and active relationship with what Impett infers as a "computer stored score" (Ibid). We will see a similar example where instrumental performance influences score in the work presented in this thesis, but from the perspective of the composer-technologist. The hardware and interface for the *meta-trumpet* were developed by technologists at STEIM, Amsterdam.

Luciano Berio's (1925-2003) work *Cronaca del Luogo* (1999) employed the EyesWeb system developed by Camurri et al. (2000). In an effort to interpret and understand human movement, "Rudolf Laban's theory of effort (1963)" (Camurri et al., 2000:60), was explored in order to "develop a modular system for the real-time analysis of body movement and gesture." This resulted in an "open hardware platform that incorporates different sensor systems...including...two video cameras, and wireless body sensors"

(Camurri et al., 2000:57). Not only applicable to the music domain, the EyesWeb system has been used by thousands of users worldwide for scientific research, education, and industry applications, and continues to be an important device/system, that demonstrates the exciting possibilities when composers have the opportunity to work with highly specialised experts (technologists) in the chosen or necessary technological field.

Interactive composition can also lead to the dissemination of complex instruments into various components, that themselves become agents of synthesis. In Juraj Kojis's (1976) *Three Movements* (2004) for piano and live electronics processing on a computer running MAX/MSP, a fascinating approach is shown, whereby the piano is extended by its various parts being used as individual components of that extension. For example, as described by Kojis (2005), the opening movement of the work makes use of the piano case as a percussion instrument. With a contact microphone attached, any attacks on the divider inside the piano results in a sound that "resembles a pulsating didgeridoo" (Ibid). Furthermore, according to Kojis, the resulting combination of the processed sounds and MAX/MSP patch filters "completes a digitally extended case instrument with new resonant characteristics" (Ibid). This is an important example of how extending any instrument can essentially create 'new instruments' as a result of the hyperinstrument. The system created for the composition was developed by an independent MAX/MSP expert to realise the required patcher objects.

The EyesWeb system was again employed in composer Roberto Doati's (1953) piece *L'apparizione di tre rughe* (2004), to track the coloured fingers of a guitarist's left-hand. Whilst the modular system of the EyesWeb package were used, further extensions to the system were added in the form of MAX/MSP patches to "apply live digital effects to the captured sounds" (Burns, 2007:20).

The final example, The Multimodal Music Stand (MMMS) as described, developed and tested by Overholt et al., (2009:69), uses "e-field sensing, audio analysis and computer vision." This system "captures a performer's continuous expressive gestures and robustly identifies discrete cues in a musical performance" (Overholt et al., 2009:69). The composer JoAnn Kuchera-Morin (1951) wrote *timeandagain* (2009) as a test-bed work for the MMMS. This is a superb example of the collaborative nature between composer and technologist(s).

2.2.1.3 - A Summary

As depicted in this brief historical overview of HCI applications in the musical domain, particularly where used by composers, the world has witnessed some interesting innovations to solve creative problems. Herbert Brün (1971:14), was very much aware of the important role the composer can potentially have in determining technologist output and vice versa, and was in principle adopted as a kind of *lingua franca* for this thesis: "The

construction of models for problem-solving in the broadest and most general sense is the goal which Technology and Composition have in common." This is *the* essence of the composer-technologist paradigm.

This history also reveals that the HCI community has invented a large array of devices, ranging from the mechanical (such as data gloves, pens, wands etc.), to those that perform audio signal processing, bio-sensing, video recognition and so on. These can however often be obstructive (Burns, 2007:22), particularly to musicians when they are haptically connected. However they may offer some form of theatricality, which can be important in performance (Emmerson, 2007:137). Non-haptic devices, whilst less obtrusive, are however not without their problems, and there is still a lot of work required in this area to perfect these systems. Some of these problems will be revealed in Chapter 5 that deals with the video recognition heuristic developed for the compositional experiment pertaining to this investigation.

2.3 - Introduction: Technology as Heuristic and as the Composer's Tool

Thus far, critical examples of HCI in the musical domain, technology and heuristics have been contextualised with regards to compositional process within the latitude of this research. The technological systems of both hardware and software, have been defined in Chapter 1 as various techniques or tools embraced to act as heuristics. The finer details of these are presented in the section that follows. Thus, from this point on,

it is possible to view *technology* as the collective term for all these concepts, as a kind of qualifier to the overall individual items of *technology* within the context as presented. Discussion now resumes with this applied specifically to the process of composition, where the mechanisms of each are viewed as co-dependent parts of the qualifier.

2.3.1 - Tools Identified

Where compositional ideation precipitates the use of technology, certain tools or 'techniques' as generally described by Becker & Eckel (1994), are necessary. Detailed in sections 4.1, 4.3 of Chapter 1, and from the analysis of the original compositional idea and research questions, the necessity to anticipate the testing, use and development of a non-haptic detection system to track the fingers of a violinist was revealed. This was undertaken within the scope of finding a cost-effective heuristic to those problems.

Cost-effective can generally be assumed to contain aspects that are free. In the computing world, free applications and platforms are usually open-source and fall under the BSD-license²⁴ and/or GNU-license²⁵ type and criteria. This was an important consideration since there was no pre-determined budget for the project. At the outset, it was assumed, that the

²⁴ Refer to the definition here: <http://www.linfo.org/bsdlicense.html> (accessed 2014, November 14) for more information.

²⁵ Refer to <http://www.gnu.org/copyleft/gpl.html> for information on this license type (accessed 2014, November 15).

setup and acquisition of a camera system and necessary peripherals would require funding, and therefore any other systems would need to be low-cost or indeed free. The identification process is divided up into two parts: software and hardware.

2.3.1.1 - Software

Ostensibly, the category of software may also be divided into two broad sections, namely those of (a.) video and (b.) audio editing and rendering tools. What became apparent in the early investigatory stage of this research undertaking was the fact that some programs can perform both categories simultaneously. Thus, the choice of software centered on its ability to be adaptable and accessible, whilst able to stream live video feed and process any information from the feed in real-time. For these reasons, the well-known and previously mentioned practical object-orientated and open-source program PD, was chosen. Version 0.42.5 was used at the time of this research project. Section 1.3 of Chapter 5 deals with PD in more detail. Puckette (2007:15) describes the program as "an environment for quickly realizing (sic) computer music applications, primarily intended for live music performances." The program is a powerful tool able to manipulate audio through timed-events, stream audio data off hard disc and also synthesise sounds using an array of synthesis techniques.

Additionally, of paramount importance is the program's feature allowing simultaneous video feed manipulation and control. The library that functions as the tool to convert image matrices in PD is called Graphics Environment for Multimedia (GEM)²⁶. This is maintained and developed at the Institute of Electronic Music and Acoustics (IEM)²⁷. Overall, the choice to use PD was however not free from its challenges, nor straightforward with regards to the acquisition of knowledge and experience, which was rather precipitous. For example, through observation of the literature available online about GEM and discussions on the forums, it was revealed that running simultaneous audio and image processing in one session of PD is not advisable, since the way in which processor core threading is done will cause audio stutter. This was an important early discovery since it would mean more hardware acquisitions in the system development phase, as well as learning how to synchronise two separate instances of PD, each dedicated to either (a.) video and (b.) audio processing.

Besides the technical issues uncovered, it was necessary to learn how to program the patches through PD's language and format. This was taken on with no prior experience or use of the application. This is however obviously advantageous when committing to undertaking the compositional process within the composer-technologist paradigm.

²⁶ For further information see <http://gem.iem.at> (accessed 2014, November 15).

²⁷ For further information see <http://iem.kug.ac.at/en/institute-of-electronic-music-and-acoustics.html> (accessed 2014, November 15).

With the choice for the fundamental core program made, it was necessary to investigate externals that could work with PD where the processing and recognition of image data could take place. Since the principal task to perform was the recognition and tracking of finger-tips of a violinist's left-hand, the Open Computer Vision (OpenCV)²⁸ library version 2.4.6.1 was available at the time and was therefore chosen. There are several reasons for this:

- firstly, it is open-source and free, in-line with the considerations outlined earlier;
- secondly, it contains a feature for object recognition and extraction that makes use of a powerful set of tools in the library which can be trained through a process of machine-learning;
- and thirdly, there exists a binding of the libraries for PD which allows for integration into the core-program.

The work of Viola & Jones (2001) together with Gómez & Degoyon (2009), was investigated to determine this, particularly where the discussion of the possibilities utilising machine-learning revealed an "approach for visual object detection which is capable of processing images extremely rapidly and achieving high detection rates" (Viola & Jones, 2001:511), and where "Computer Vision is acquiring a growing relevance in the field of interactive arts" (Gómez & Degoyon, 2009:1). Whilst, Viola & Jones (2001)

²⁸ For further information see <http://opencv.org> (accessed 2014, November 15).

explored the creation of face-recognition systems, they rely on the principal and technique of object identification. Overholt et al., (2009:71), describe the use of the Viola & Jones (2001) OpenCV method to detect the gesture of a flautist's head to control cueing events in a musical work. Whereas this application of the OpenCV method is not directly related to instrumental gestures, a similar method as created by Viola & Jones (2001) was chosen for training the recognition of finger-tips. Section 3 of Chapter 5 deals with this in detail.

The OpenCV community world-wide relies on the creation and availability of trained classifier and image data-sets to perform various applications of the system. It should be noted that whilst many thousands of data-sets are available that cover everything from car, face, building, bicycle and animal recognition etc., (Dawson-Howe, 2014:160-161), there appears to be no trained classifier specific to finger-tip detection²⁹.

Several important idiosyncrasies of the OpenCV detection system were gleaned from the overall investigation into the use of the software.

They are:

- the system must run with a minimum of 15 frames per second rate;
- the window size of the incoming image stream must be no less than 384 by 288 pixels (Viola & Jones, 2001:511);

²⁹ This was determined using standard search methods online, such as the search prompts available on the data-set websites, and using Internet search engines such as Google.

- ideally, the computer system should contain an Intel-compatible processor as the software makes use of optimised assembly routines, such as the Intel Integrated Performance Primitives library (IPP) (Gómez & Degoyon, 2009:1);
- training and processing time of the machine-learning stage can be lengthy (several weeks, with many attempts), depending on the data-set size of prepared images being used;
- several 3rd party applications and tools will be needed to begin the data-set preparation, such as an Integrated Development Environment (IDE) for compiling a program in C++ to allow for the ROI (Region Of Interest) of the positive images to be determined as well as image editing tools (Seo, 2007).

Now that the software particulars have been outlined, the hardware that was researched, obtained and then used is discussed.

2.3.1.2 - Hardware

The computing systems chosen and used for this research project were Apple's Mac Mini i5 desktop computers and peripherals (two were acquired to allow for independent audio and video processing), each running Mac OS X version 10.9.5.

Since the system requires the live streaming of video feed, several options that utilise communication via USB (Universal Serial Bus) and

fire-wire where investigated. The market for such devices is enormous, and it was completely necessary to narrow the scope of this down, by limiting their inherent applications and system compatibilities. It was determined that:

- the device should be plug-and-play and preferably natively supported by the Mac OS X operating system without the need for drivers or independent software;
- it should be capable of live streaming;
- it should have a frame rate of no less than 15 frames per second;
- it must be capable of processing a minimum of 384 x 288 pixels (from earlier).

The need for these explicit specifications, narrowed the search down to devices manufactured by a company called Unibrain^{®30}. Initially, the entry-level Unibrain[®] Fire-i^{™31} device was selected and acquired for the purposes of testing and eventual development of the system. Incidentally, this same camera was used in an application of a table-top tangible user interface (TUI) for designing a system to use space as a musical parameter and instrument (Gasteiger, 2010:1, 60). As more funding became available in the very late stages of the research project, it was possible to acquire a

³⁰ See <http://www.unibrain.com> for more details (accessed 2014, November 15).

³¹ See <http://www.unibrain.com/products/fire-i-digital-camera/> for more details (accessed 2014, November 15).

more robust and accurate camera. The Unibrain® Fire-I 785c³² camera proved the most promising in its specifications, where it offers much higher resolution and picture size (up to 1280 x 960), with the option to add any focal-length C-mount lens to the camera. Whilst the camera is natively supported in Mac OS X and works in the operating system's bundled video products with plug-and-play ease, it unfortunately causes PD to crash each time a video render is activated. This was immediately reported to the GEM developers, and so far a response has not been forthcoming. However, the product was tested in PD's sister commercial program MAX/MSP and it functions as expected. This means, that the project would have to be ported to MAX/MSP in the long-term if the developers at IEM do not intend to rectify the bug.

2.4 - The Next Step

Viewed through the lens of both compositional ideation and process as defined, and where the composer (author) assumes the role of a technologist, a discussion now ensues where the link between the kinetics of a musician and the system as theorised and created are explored. This is achieved firstly with a discussion of musical gesture, followed by a contextualisation of psycho-acoustics that relate to the musical gestures studied, and then finally how these concepts were applied to the creation of *In-Gest*.

³² See <http://www.unibrain.com/products/firewire-800-industrial-cameras/#details2> for more details (accessed 2014, November 16).

CHAPTER 3 - GESTURE

Music is tending towards spatialisation, towards the exclusion of sequential presentation, towards objectification and reification, and it brings elements from the world of the eye into that of the ear.

- Robert Beyer, *Das Problem der "kommenden Musik"*, 1928

3.1 - The Study of Musical Gesture

The study of gesture in music, i.e. gesture that relates to or occurs from musical creation, performance and reception, has received a thorough and detailed investigation in recent years. This is particularly evident from the significant body of literature available on the subject. The detailed contributions of Wanderley & Battier, Eds. (2000), Hatten (2004), Gritten & King, Eds. (2006 & 2011), Leman (2008), Godøy & Leman, Eds. (2010), and many more, support this. The approach has generally been rooted between the semiotics of language and music, where Schneider³³ (2010:69) highlights the "linguist Manfred Bierwisch('s) (1979)" discussion of the two, where "music in general bears a 'gestural form' (*gestische Form*), whereas

³³ The passages in English that occur in the citation are as translated by Schneider.

language basically exhibits a logical form." This is demonstrated by the writers of language "to be able to 'say' (*sagen*) something", whereas it is "essential for a composer or musician to be able, by means of music, to 'show' or 'demonstrate' (*zeigen*) something" (Ibid).

Gesture in musical context is considered as movement that results from specific circumstances, containing distinct connotations (Hatten, 2001:L1). This is viewed by Hatten (Ibid), as a "holistic process", where the separate components of musical information such as "melody, harmony, rhythm and meter, tempo and rubato, articulation, dynamics, and phrasing", are parts of the united whole. These are all elements of musical form which are represented by features of the Western notation system, which when used together create musical structure. Traditionally, the study of musical gestures has focused on the analysis of musical notation and its ability to convey meanings that refer to or afford gestural contexts. Critically these principles and mechanisms relate to the questions posed as part of the original compositional ideation in Chapter 1. Kühl (2011:123) elucidates: "The most important, stable element in a musical semantics is the primary signification from musical phrase to gesture and from musical gesture to emotional content and social belongingness." The mechanism that drives this transference of meaning is usually initially notation, and thus musical notation when occurring in this context may be viewed in musicology as a form of semiosis – a signifier that conveys information. Martin (2014:189), views this as "semiotic resource(s) of

music”, and additionally viewed as a process of cognition³⁴ between motor systems and musical elements, forming a musical semiosis (Kühl, 2011:125). Moreover, at its base-level, notation is a set of instructions, an “instruction manual” as Read (1979:24) puts it, with each part of the instruction being small symbols that act as signs (containers of more information than just the symbol). Each of these signs relate to each other in a contextual matrix that now has the possibility to convey some meaning (semantics). However, Hatten argues that notation alone cannot fully embody the meaning contained by musical gesture:

“Given the importance of gesture to interpretation, why do we not have a comprehensive theory of gesture in music? Historically, one reason may be that musical notation, which is largely digital or discrete in its symbols, cannot easily represent the continuities of gesture” (Hatten, 2001:L1).

This is further supported by Rink et al., (2011:267), where the characteristics of music cannot be “fully encoded within musical notation.” Nevertheless, the study undertaken here attempts to demonstrate that with the possibilities afforded musically in electroacoustic music, a better understanding of gesture and its relationship to score (notation) is possible. However, no attempt is made to decode meaning in notational semiotics through gesture analysis, as this is beyond the scope of the

³⁴ Music *cognition* (cognitive processes) is explored in more detail in Chapter 4.

thesis. Instead, clarity is provided to this process by comparing the structures of the compositional elements in the acoustic score to those of the electroacoustic score and how, through the creation of movement, the two relate to or afford and support each other with the aid of a gesture heuristic in the form of a finger-tip "action tracker" (Emmerson, 2007:136). An in-depth account of compositional elements is undertaken in Chapters 5 and 6.

Where music creates movement in the form of gestures where live instruments are used, they are visible³⁵. Therefore, these visible motions are the containers of information. However, by this context and meaning of gesture, not just any movement qualifies as a gesture, and they need careful description and classification in order to be meaningful within the context of a study of gestures in music. Hatten (2001:L1) qualifies this by stating, "Whenever characterization (sic) is involved, we can speak of meaningful gestures." These are referred to as "structured interactions" (Leman & Godøy, 2010:9). Previously published research on gesture categories has presented the domain with various well-established terms to describe this particular branch of gestural study. Cadoz (1988:5) has termed them "instrumental gestures" as previously³⁶ shown. Similarly, in citing Francois Delalande, Cadoz & Wanderley (2000:39) reflect on

³⁵ In acousmatic music this is not the case. The notion of 'image in the mind' associated with compositional ideation is discussed later in this chapter. However, the reader is directed to Barreiro (2010), for more information that concerns listening modes of acousmatic music and gesture.

³⁶ See section 2.5 of Chapter 1.

“effective gestures”, whereas Jensenius et al., (2010:23) refer to the same gestural category as “sound-producing”.

Musical gesture on the other-hand, refers to a compositional element, phrase or sound object that conveys an abstracted meaning. This is because the work here is viewing *musical gesture* through the lens of composition and not performance and reception as in Godøy & Leman (2010). To support this, Hatten (2001:L2), further argues that gestural articulation as part of compositional form and structure has “consequences for our understanding of style and style change as well as for the interpretation of a work.” This is critical in the task of performing musical analysis. Until fairly recently however, this was not fully possible without particular advances in certain technologies, such as “motion capture” and software to proceed with examination of captured data, which allows researchers to study movement from music with extremely high precision (Leman & Godøy, 2010:3).

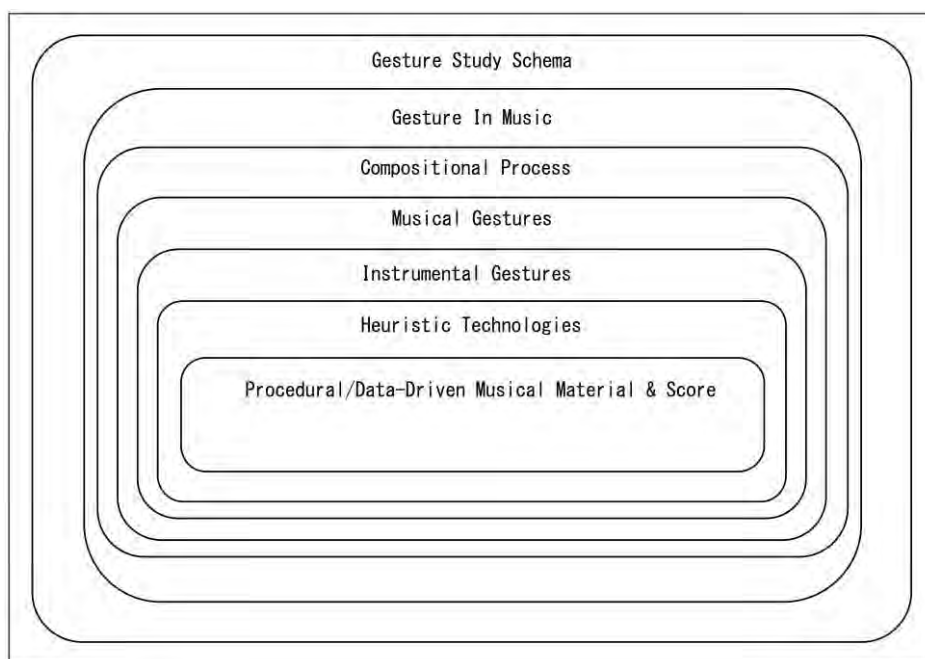
In further consideration of the context of the study of gestures relating to the work *In-Gest* as presented in this thesis, an attempt is pursued to uncover how a process of writing musical gesture creates the possibility of targeted instrumental gestures, which can be used to create a partnered electroacoustic score (section 4 of Chapter 1 outlines the research approach and questions relating to gesture in the compositional process).

In essence, there are two dynamics present in the study of gesture in this thesis. They are:

- the study of technology as a heuristic tool to compositional idea where instrumental gesture plays a role in the formation of the technological heuristic, and;
- the study of the embodied meaning (compositionally) of musical phrases through their gestural articulation and how this relates to compositional process – i.e. can a fully realised semiosis of musical gesture become a compositional parameter?

The figure below illustrates the overall process of gestural study as approached:

Figure 2. – Gesture Study Schema



As the schema contains a process that generally relies on gesture as part of its integral formation, 'gesture in music' as an overriding container is pertinent, since as previously discussed, music can be seen as a 'gestural form'. Interestingly, this method of study outlines an arrangement that appears tiered, where a perceived linear series of events as parts of a greater process, interact with each other to produce the final result. Evident from this, is a duality occurring between the composer as instrumentalist, where computer programming that produces sound is the 'instrumentalist' and, that of the composer as creator of structures using that instrument. The composer thus assumes the dualist role of both instrumentalist and structuralist. Cadoz (1988:4), places this within a "materialised model" for the organisation of compositional structures which describes the relationship as "hierarchical...exist(ing) between the production of material and its organisation." This is only possible where gesture is used as part of a compositional process, because without capturing the information emanating as musical movement (instrumental gesture) from musical gesture and using it as part of the composition, there is no hierarchy. The organisation of the study of gesture as proposed in this schema approaches the phenomenological method of gestural analysis as proposed in section 2.5 of Chapter 1.

3.2 - Musical Gesture and Meaning

In this section of the exploration of gesture in music, the focus turns to a deeper look at the gestures themselves and their properties with respect to their structures, functions and affordances. This is undertaken within the framework as outlined in section 2.5 of Chapter 1, where descriptive analysis (phenomenological) and functional approaches are applied (Gadoz & Wanderley, 2000:33). Analysis from the performer's point-of-view is undertaken in Chapter 5 once several sessions of gestural tracking have been explored with the heuristic.

3.2.1 - Structural Outlines of Musical Gesture

As the 'Gesture Schema' previously outlines, there is a linear process in the methodology adopted in the study of gestures pertinent to this investigation. The stages are individual, yet linked. As discussed in some detail so far, compositional process - the amalgamation of ideation and the act of transforming the idea into material - has elements that form semantics in the form of musical gestures. Leman (1999:115) supports this where musical "perception is largely based on auditory image representation and gesture representation rather than attribute representations."

The division and outlines that follow work towards a study of “gesture space” as researched by Jensenius et al., (2010:20). In consideration of the compositional elements referred to as “sonic gestures” (Hatten, 2006:1), or generally as “musical gestures” (Godøy & Leman, 2010:3), and their position in the Gesture Study Schema (Figure 2.), the details below place compositional ideation, process and musical gesture in that context.

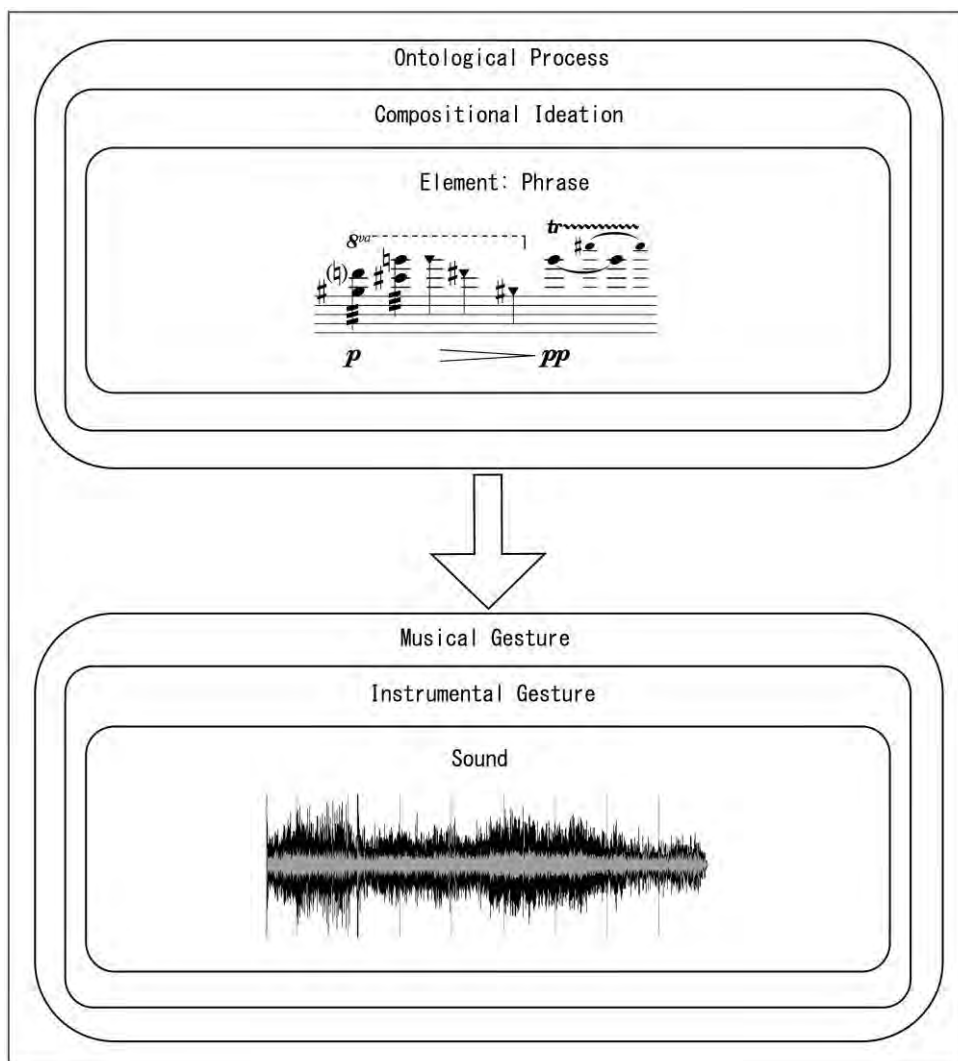
3.2.1.1 - Gesture Space and Musical Gestures: A Contextual Structural Outline Pertaining to Compositional Process

The outlines of compositional process in section 4 of Chapter 1, as well as the descriptions of ontological processes giving rise to compositional idea in section 1 of Chapter 2, refers. Kühl (2011:127) argues for a signification of cognitive processes through a semiotic practice, where “phrase” creates “sound pattern” simultaneously with “musical gesture”. In a similar disposition, the *Musical Gesture Signification Model A* proposed in Figure 3., overleaf, represents the methodology for signification as undertaken so far.

Extracted from the “Float” section of *In-Gest* (See Appendix I for complete score), the example illustrates the stages in creation of gestures in the musical space from the initial conception of the phrase to what is heard. Apropos the details of the work in Chapter 6, each section of the work has been conceived as separate gestural images, evoking different intentions,

affordances, feelings or sensations. For example, walking, floating, ascending, descending, dancing and so on. The intention that “*Images of gestures seem to be efficient in evoking images of sounds*” (Godøy, 2004:56), is of paramount significance here, and explored where the electroacoustic score supports these gestural affordances.

Figure 3. – Musical Gesture Signification Model A

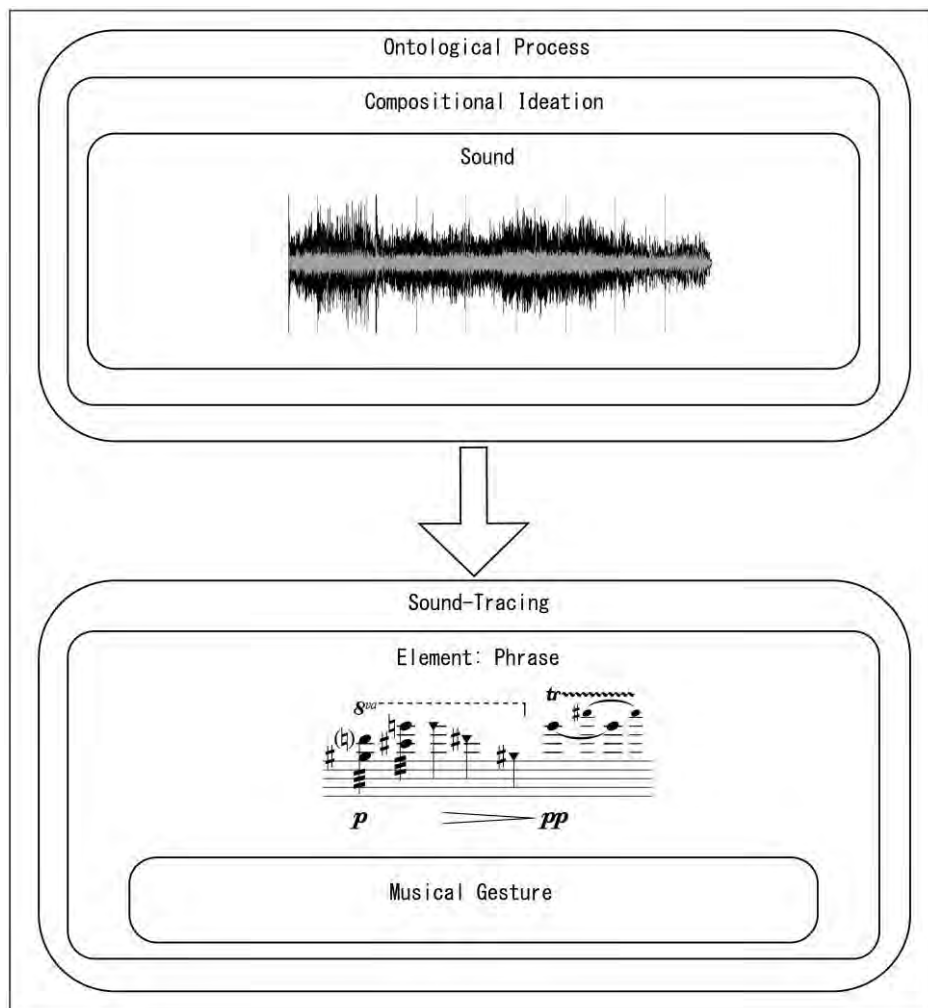


The process in model A however, is only one possibility. Experience in composition and process reveals the likelihood of alternative routes

within this model's structure. Instead, as Godøy (2010:103) points out "musical sound has great power to make us move, or to create sensations of movements in our minds." Of particular interest here is the last section of this statement 'sensations of movements in our minds' that often takes the form of mental images as part of initial ontological processes during compositional ideation. The composer 'hears' the movement cognitively and then outlines this movement in the form of a musical gesture. This is referred to as "sound-tracing" (Godøy, 2010:117). Whilst this sound-tracing process usually describes auditory perception from an external stimulus i.e. listening to sounds, the sound-tracing here is an internal process, occurring without any external auditory stimulus. As part of the explorative research undertaken in Chapter 4, the cognitive processes responsible for this internal process are investigated. If the composer chooses to audition gestural affordances by actively listening to material and then tracing the gestures cognitively, the process is analogous to sound-tracing by external stimulus. Figure 4. - *Musical Gesture Signification Model B* overleaf, outlines the process of 'sound-tracing' where it occurs as part of compositional ideation.

This model proposes a form of integration of the practice in most Western music traditions where usually the score is separated from performance (Godøy, 2010:109). The composer is imagining performance in terms of gestural structures and the score is realised in tandem with this.

Figure 4. – Musical Gesture Signification Model B



Furthermore, Godøy (2010:110) states:

“Western musical thinking often tends to ignore the fact that any sonic event is actually included in a sound-producing³⁷ gesture, a gesture that starts before and often ends after, the sonic event of any single tone or group of tones.”

³⁷ See section 2.1.1.1 of this chapter for details.

Additionally this implies, a "*motormimetic*³⁸ element in the perception and cognition of musical sound" (Ibid). Therefore, as a result of viewing music from this position, it is possible to apply the study of gesture in "schemas" as proposed here in the creation and perception of musical sound (Ibid). Correspondingly the choice to undertake a project where a gesture heuristic focuses on small gestural units (movements of finger-tips for example), is supported by the application of studying gestures as part of a 'schema' as outlined. Furthermore, the application of these models within the schema developed places this part of the analysis in the "phenomenological" and "functional" approaches as intimated by Cadoz & Wanderley (2000:33), discussed in section 2.5 of Chapter 1. An examination now commences pertaining to the instrumental gestures, which themselves are broken into *gestural units*.

3.2.1.1.1 - Gestural Units

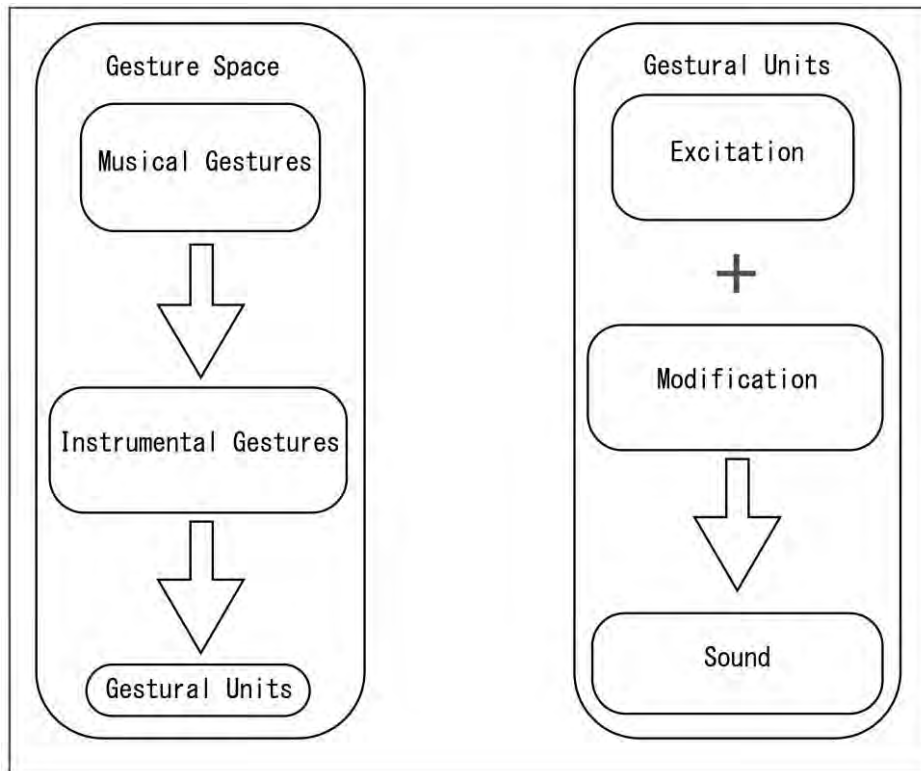
Research undertaken to identify aspects of instrumental gesture has resulted in the dissemination of many smaller parts of the larger category. The movements vary from basic haptic functions such as striking a key or stopping a string node to movements that communicate with other instrumentalists or the audience or even movements that are supplementary to or accompany the sound-yielding gestures (MacRitchie et al., 2013:87). Respectively, Jensenius et al., (2010:23), refer to these other forms as

³⁸ "Motor-mimesis" is the research field that aims to explore understanding of musical sound through "mental images of sound-associated actions" (Godøy, 2004:55).

“sound-producing”, “communicative” and “sound-accompanying”. Not yet mentioned are “sound-facilitating” gestures, which support the “sound-producing” gestures in particular ways, such as a trained arm movement to create a particular timbre of sound (Jensenius et al., 2010:24). In Figure 5., *Gesture Space* on the following page, the various elements of gesture in the music domain are outlined, whereas *Gestural Units* details the properties of the sound-producing gestural unit.

Since the sound-producing instrumental gestural unit is of primary concern here, the dissemination and detailed study of its intrinsic construction is necessary. This is because producing sound with movement is a complex process of multiple functions each with their own purpose such as “excitation” and “modification” (Jensenius et al., 2010:24). ‘Modification’ referring to changing timbre or loudness, such as vibrato, pitch-bending or bow-force pressure changes on stringed instruments, whereas ‘excitation’ refers to creating the initial sound, such as pressing a key or stopping a string.

Figure 5. – Gesture Space and Gesture Units



The *Gestural Units* portion of Figure 5., outlines how these smaller parts fit together to form a gestural unit. Since the heuristic developed for the purposes of this study deals with the tracking of finger-tips on a violinist’s left hand, excitation gestural units are explored, but modification gestural units are ignored as the system developed is not concerned with nor designed to capture modification data. Besides the choice to concentrate on the tracking of finger-tips (an under researched area as discussed) and not bowing, one of the particularly influential factors of gesture capturing that informed this decision is what is referred to as the “degrees of freedom problem” (Dahl et al., 2010:37). The nature of human movement results in the possibility for a high

variance in the performance of the movements, particularly in an interpretative framework such as a musical work. As it is, it proved essential to implement programming strategies³⁹ in the processing of the captured gestural data in the heuristic developed to curtail this problem and its implications with regards to sound-producing gestures.

3.2.1.2 - Functionality

Sound is heard when air molecules are caused to vibrate through an object's stimulation owing to the principal of "resonance" (Halmrast et al., 2010:185). Since excitation gestural units are explicit in the production of sound, the excitation unit when referenced in the bowed stringed instrumental domain has two parts, namely (1.) stopping and (2.) bowing, plucking or striking. This is so since the creation of sustained sound relies on the simultaneous occurrence of both parts unless an open string (where no stopping is required) is called for. The pitch is determined by the length of the string being excited, which is achieved by the stopping portion of the gestural unit.

In terms of its function related to the gestural unit of excitation, the heuristic developed has two main purposes:

- locate finger-tips (objects) on the left-hand of a violinist;

³⁹ See sections 2.1.2 of Chapter 5 and 2.3.2 of Chapter 6.

- send the (x,y,z) co-ordinates of the located objects and their associated identifier to the system for processing.

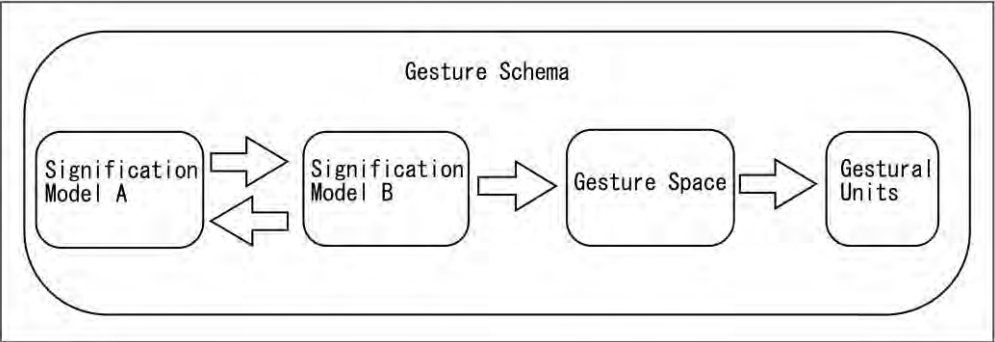
Thus finally, the heuristic is concerned with the stopping action portion of the excitation gestural unit and its associated function in sound production. Jensenius et al., (2010:25) refer to the overall functionality of gestures pertaining to a specific situation as "dimension spaces". These differ according to presentation, such that a violinist's gestural space would contain differing units of function (excitation, modification etc.), than say would a dancer whose space would contain more sound accompanying gestures. As described previously, the analysis of the gesture space where the units are broken down into their smaller components, approaches the "intrinsic - from the musician's point-of-view" (Cadoz & Wanderley, 2000:33).

3.2.2 - Summary

In conclusion, the dissemination of a chain of events in the perception and reception of musical gesture has taken place within the analytical approaches as delineated in section 2.5 of Chapter 1. The overall approach included an application of a schema-based study with the presentation of various models of gestural processes relating to composition. The figure on the following page summarises these models and processes.

Furthermore, this has been undertaken within the overall view of the development of a gesture heuristic that uses the knowledge gained from the study to solve a compositional problem as defined and contextualised. In essence, the study has explored, as undertaken through the lens of compositional process, what is a hermeneutical cycle, (Godøy (2010:119) refers to "hermeneutic circle", perhaps in reference to phenomenology as that described by Ramberg & Gjesdal (2014)), between composition and performance of musical gestures and the resulting instrumental gestures from the interpretation of a musical score.

Figure 6. - Complete Gesture Schema



In summation, even though this thesis explores gestures that emanate from the articulations caused by the music, it does not as seen undertake to include an overview of the meaning of the musician’s gestures. It is primarily concerned with the effect of thinking of gesture as a compositional parameter – or the effect of gesture on compositional process where a technological heuristic is used as a tool to achieve said process. Thus the communication of the embodied meaning, through gesture,

is a result of this hermeneutic relationship between musical phrase and the resultant gesture. Noteworthy is the often used application of motion capture systems and gesture study models to analyse the recordings of human performance in order to study musician expressions (Dahl & Friberg, 2007:434). In contrast, the study presented here captures the gesture movements to ascertain the direct interaction between the composition of musical phrases and their subsequent creation of instrumental gesture during performance. The stage is now set to explore music cognition and how the completed study of gesture in context relates to that.

CHAPTER 4 - MUSIC AND MEDIATION TECHNOLOGIES

What we have to learn is what the human brain and ear thinks is beautiful.

What do we love about music? What about the acoustic sounds, rhythms and harmony do we love? When we find that out it will be easy to make music with a computer.

- Max Matthews, 2011

4.1 - Music Cognition and Psychoacoustics

The discipline of cognitive musicology is about 40 years old now. The essence of its focus has shifted from the view of a Cartesian centered approach where the mind is completely separate from the body to that of a more "naturalistic approach" concerned with the influence of the physical environment (Leman, 1999:93). As described in section 2.8 of Chapter 1, *psychoacoustics* is concerned with the physical properties of sound and their subsequent cognition. Cognition of sound is thus a sub-category of the study field of psychoacoustics, and indeed the material presented in Cook (1999) and Leman (2008:29) stands testament to this. Furthermore, the disciplines are "complementary", where "representational levels can be

arranged hierarchically so that acoustic features give rise to musical concepts" (Laden & Keefe, 1989:12). As seen in section 2.8 of Chapter 1, psychoacoustics had laid the foundations for the soon to be dealt with discipline of *music cognition*.

As also discussed in the previous chapter, there are several cognitive/ontological processes taking place in the unfolding of composition, and where a study of musical gesture as a phenomenon undertook to reveal the processes of both internal and external perception and reception of sound. Whilst important principles have been contextualised and discussed as part of ontological processes, and therefore hint at cognition, the mechanisms of cognition that embody musical thoughts when undertaken in the framework of a study of gesture heuristics have not yet been viewed. This is so, since the use of the body to explore the process and product of composition, itself a cognitive phenomenon, is an indication of the embodiment of cognition - all possible within a study of gesture.

As we have seen, gesture is movement or action where these "actions may form a link between the mental and physical worlds" (Leman, 2008:14). In the realm of original composition, cognition is thus an important consideration when undertaking explorations of the processes in the perception of sound and music. Furthermore, by virtue of the fact that human beings experience all occurrences, whether musical or otherwise, as

procedures that involves "interpretation in the brain" (Shepard, 1999:21), the inclusion of cognition in a study of compositional process cannot be overlooked.

Establishing the mechanisms of auditory cognition in terms of imagery and using the information to study compositional process is problematic. This is as a result of the way in which humans have evolved to process auditory information. We are wired to perceive the world in terms of images, but use hearing to add information to that perception (Pierce, 1999:89). This infers that our external auditory perception of the world can only complement our understanding and interpretation of our surroundings when we have sight. But composition is, as has been demonstrated with the aid of *Gestural Signification Models A and B* in the previous chapter, both an internal and external process that relates imagery to auditory information in cognition. Therefore, the world of the composer is quite different in perception when using sound to structure images that represent intentions, affordances or semantics. Consequently, part of the intention of this study is to relate the cognition of music to the creation of musical gestures as mediated by an instrumental gesture heuristic.

4.1.1 - Music and Sonic Forms

In his treatise on Embodied Cognition, Leman (2008:17,19) argues that "sonic forms" are the formal structures of music, which in themselves do

not contain meaning, but when articulated in a musical setting reveal “corporeal signification” and are therefore significant in providing semantics. The discovery of the body acting as mediator of cognitive information, especially in music, has necessitated the need to develop what are termed “mediation technologies” where sonic forms are combined with technology (Ibid). The idea is that these technologies assist in bridging the gap between “mind and matter” (Leman, 2008:49).

4.1.1.1 - Mind and Matter

The notion of ‘mind and matter’ and an exploration of the link between them is an important telos in this work. The answers to the compositional questions proposed and the techniques involved to produce them are not possible without an understanding, at least at an introductory level, of this significant dichotomy. According to Leman (Ibid), “the relationship between mind and matter is one of the main themes of the history and philosophy of music research.” This is very much evident from the anthropological, psychological, sociological, and aesthetics research endeavors undertaken by not only musicologists but also researchers in vastly disparate disciplines such as science and medicine. Whilst a complete history of this is well beyond the scope of this thesis, standing testament to this statement is the early ground-breaking treatise of Helmholtz (1895:1), whom opened his *Die Lehre von den Tonempfindungen als Physiologische Grundlage für die Theorie der Musik* with:

"The horizons of physics, philosophy, and art have of late been too widely separated, and, as a consequence, the language, the methods, and the aims of any one of these studies present a certain amount of difficulty for the student of any other of them; and possibly this is the principal cause why the problem here undertaken has not been long ago more thoroughly considered and advanced towards its solution."

Leman (2008:49) supports this where in late nineteenth century research, the "connection between subjective experience and matter has to do with the human brain" is explored. Following this, *cognitive science* emerged in the mid twentieth century as the "interdisciplinary study of mind and intelligence, philosophy, psychology, artificial intelligence, neuroscience, linguistics and anthropology" (Thagard, 2014). A short time later in the 1970s, "the youngest of the cognitive sciences", that of *cognitive musicology* emerged, which concerns itself with the "modeling of musical intelligence in its many forms" (Laske, 1988:43,44). Of critical importance, here is what Laske (1988:45) describes as the ability for computer programs to add credence to postulates about musical understanding and provide a platform for designing "structured task environments (such as programs for interactive composition)." What is intimated is that computers must form the indispensable platform and tool for undertaking and exploring the field of cognitive musicology. Indeed, compositional process has been defined within the parameter of a *task*

environment as seen in section 2.2 of Chapter 1. Not only that, but the system developed and detailed in section 2.2 of this chapter, is an 'interactive composition', and assumes the role of exploring compositional process from a cognitive perspective when combined with analysis,

Discussed now is the 'matter' part, particularly related to body movements that are used in the application and development of the mediation form presented here. Muscles and their link with the brain via nerves create body movement when activated. In research terms, this is referred to as *motor cognition*. According to Smith & Kosslyn (2006:452), motor cognition is "mental processing in which the motor system⁴⁰ draws on stored information to plan and produce our own actions, as well as to anticipate, predict, and interpret the actions of others." Furthermore, this is related to what is known as the *perception-action cycle*, which is the "transformation of perceived patterns into co-ordinated patterns of movement" (Smith & Kosslyn, 2006:453). In the musical domain, the coupling of gesture studies and the perception-action cycle broadens our understanding of what we hear and the resulting action behind it (Godøy, 2010:119). This certainly seems true when the observation of sound-producing gestures immediately anticipates the sound we hear as a result of them. In the cycle of a violinist's finger movement relating to the stopping portion of the excitation gesture unit, the motor-cognition is

⁴⁰ For a detailed understanding of *Motor Systems*, the reader is referred to the work of Rizzolatti & Luppino (2001).

“intrinsic” or “from the musician’s perspective”, applying the approach of observation delineated by Cadoz & Wanderley (2000:33).

The following section discusses the mind part of the paradigm in more detail, which is subsequently applicable within the mediation form outlined.

4.1.1.2 - Musical *Gestalts*

“Hearing in the head” is a description often associated with the natural phenomenon attributed to mental imagery (Thomas, 2014). In Gestalt psychology, the phenomenon is understood as the acuity of motion between static objects (Wertheimer, 2014:131). The relationship of this branch of psychology to cognitive science is clear when considering the already discussed sonic forms and the imagery presented as part of the study of gesture in Chapter 3. Thus Gestalt theory and psychology owes its lineage to a variety of disciplines that describe the mental processes of moving images and according to Wertheimer (2014:133), is rooted in a “holistic Weltanschauung, in musicology and in the psychology of cognition.”

In the focus of the notion of a musical Gestalt, it is necessary to re-iterate that “musical gesture is a cognitive phenomenon, emerging in the mind in response to musical priming” (Kühl, 2011:125), and further where it originates from the “generic level of perception, where it is tied to

gestalt (sic) perception, motor movement and mental imagery" (Ibid). Since Gestalts are a priori of wholes, being more than the sum of their parts (Wertheimer, 2014:133), the analogue with musical elements that make phrases, that are in turn made of smaller elements such as harmony, rhythm, meter, tempo, etc., but also contain external factors such as timbre, interpretation and context, is obvious and seems to support the idea of gestural images existing as musical Gestalts. Sections 1.2 and 2.4 of Chapter 6 reveal the application of musical Gestalts in compositional process.

4.1.2 - Information Theory

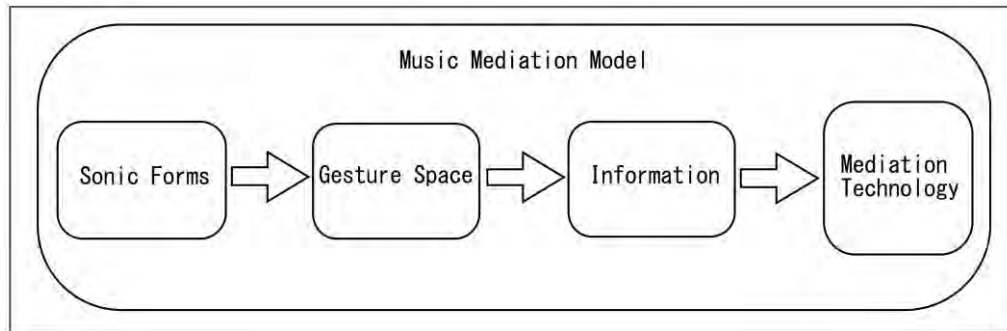
Information theory is primarily concerned with the initial stages of the process in communication of messages (Reimer, 1964:14), where early models of computational information theory dating back to the 1950s and 1960s, were used to construct meaningful objective accounts of musical analysis (Leman, 2008:30,31). Reimer, (1964:14) explains that the formation of messages can take many forms such as music, gestures, words or images, and are implicit in the intention of the choice to convey specific meanings. This is an important characteristic of information because it "provides a means of quantifying the innermost process of the musical art - the sequential development of interrelated patterns" (Reimer, 1964:16). The heuristic as developed and applied in this research endeavor is primarily concerned with using small variables of precisely those elements of

information to allow the translation of meaning of musical gestures from one to the other in the scores that accompany each other in *In-Gest*.

Computer technology deals with data that is entirely made of numbers and is based on the binary system of information, which is converted into meaningful forms such as display, audio and print-out messages. Thus using the system developed to isolate small instances of information generated by musical gestures is analogous with the early models of information theory, only in this case, the data is not analysed per se, but used as a trigger mechanism for gestural affordances. When applied in the *gesture space* as described and detailed in Chapter 3, a model for musical mediation is possible. Figure 7., overleaf, outlines the structure.

This model's structure by no means infers a generalised or preferred approach to developing mediation technologies. Instead, its form is a direct result of the procedures as researched and applied in the unfolding of the investigation into compositional process as outlined. It could however be utilised as a basis for the development of future mediation forms that are reliant on compositional process to inform their intrinsic structure.

Figure 7. - Music Mediation Model



4.2 - Mediation Technologies - Introduction

As described in section 1.1 of this chapter, the combination of sonic forms and technology creates 'mediation technologies'. The capturing of these sonic forms by gestural heuristics is of interest to mediation technologies (Leman, 2008:22). Where applied in a compositional aesthetic, Cremaschi & Giomi (2004:27) argue that:

"the most important "electronic" works of the last twenty-five years are those that have sought a mediation between the acoustic dimension and another realm - those that expanded the continuity between "electronic" sounds and "natural" sounds, enabling interaction between the different levels through reciprocal transformation."

In tandem with the review undertaken in Chapter 2 concerning both composer-technologists and composers working with technologists in

developing HCI systems for musical experimentation, composers and researchers alike have contributed to the field of HCI research significantly where technology is viewed as a mediating agent in experimental music (Hamman, 2002:114,115). Important recent developments are listed in tabular form next, particularly where they relate to string instruments and especially the violin.

4.2.1 - Recent Research and Developments

The following exemplify this attitude of experimentation by researchers in the field where gestural heuristic technologies have been developed to create mediators of musical information. All of these examples have been developed either by fully versed technologists in their own right or researchers embarking on projects with the aid of expert technologists relative to the fields required. This has resulted in complex engineering projects that have been applied in the musical domain, but are not directly related to composition as explored in Chapter 2. The intention is to place the heuristic developed by an untrained composer becoming a technologist (the author) in the context of engineering-oriented projects. Whilst these details are necessarily brief and confined to a table (see Table 1., overleaf), the reader is encouraged to investigate further by following the references.

Table 1. – Gesture Mediation Technologies

Date Developed	Description/Outline and Purpose	Application	Reference
2004	With the use of IRCAM's <i>Augmented Violin</i> , an approach for the analysis of violin bow strokes was developed. The technology was inspired in part by the <i>Hyperinstruments</i> ⁴¹ project developed at MIT.	Physical model synthesis and direct gesture interpretation using accelerometers and sensors.	(Rasamimanana, 2004)
2007	The focus of this project was to investigate the potential of corporeal measurements of violin bowing dynamics to update musicological understanding of bowed string performance and performer skills.	Bowed string physical model using advanced version of <i>Hyperbow</i> ⁴² .	(Young, 2007)
2009	The research undertaken in this project created models for analysis and synthesis of instrumental gestures on the violin, particularly related to bowing.	Commercial electro-magnetic field measurement instrument placed in the bow.	(Maestre, 2009)

⁴¹ As discussed in section 2.1.1 of Chapter 2.

⁴² Refer to Young (2007:38) for details.

Date Developed	Description/Outline and Purpose	Application	Reference
2010	Besides measuring bow force and action, pressure sensors also applied to neck to sense left hand finger pressure.	Goniometers and pressure sensors applied to bow and neck parts of the violin.	(Grosshauser, et al., 2010)

4.2.1.1 - Selected Examples

With these brief examples, it is obvious that the focus of complex engineering-based approaches to gestural heuristics in the string domain focus on the right-hand and particularly bowing actions as part of the excitation and modification parameters of instrumental gesture. What is lacking though, are efforts to track and create mediation using the left-hand as a source of information (Grosshauser, et al., 2010:273). The study here contributes to overcoming this lack of research particularly where undertaken by a composer adopting technologist skills and knowledge. What is clear, and supported by Camurri & Moeslund (2010:241), is that “music production based on visual gesture recognition can be considered as a special case of human computer interaction.”

4.2.2 - Towards a Mediation Technology Form for *In-Gest*

The model proposed earlier in Figure 7., outlines the processes taking place in the perception and reception phases of compositional process. The goal of the mediation technology in this model is to address the possibility of linking gestural affordances between two parallel scores - that of the acoustic score and that of the electroacoustic score in the work *In-Gest* as described so far.

As intimated within the current research paradigm, the use of finger trackers for information capturing of the left-hand of string players is under-developed. Particularly noteworthy is the fact that the use of video streaming technology to perform these tasks is almost non-existent where cost-effective means to achieve object-tracking and recognition is concerned. This is garnered from the outlines of the HCI technologies in Chapter 2 and the research for gestural mediation as presented in Table 1.

The goal of the mediation technology developed here is thus three-pronged:

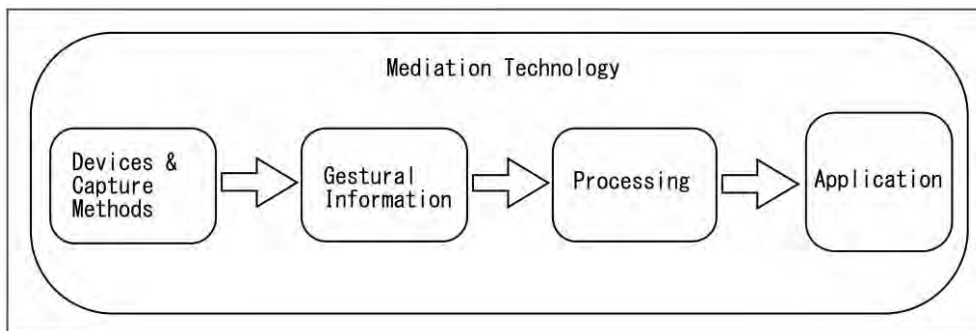
- use the mediation technology to solve a compositional problem and generate procedure and information for an autoethnographic⁴³ study of compositional process;

⁴³ See section 2.1.1 of Chapter 2.

- provide research and solutions to the under investigated paradigm of string player left-hand gestural interfaces;
- contribute to non-haptic object tracking research and provide useable classifiers for future research.

The mediation technology's structure, properties and methodology are discussed in detail in Chapter 5. However, an outline below is provided to contextualise the methodology.

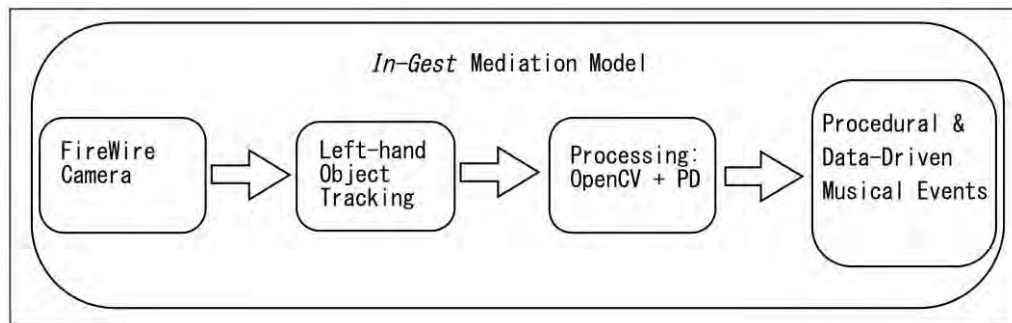
Figure 8. – Mediation Technology Methodology



In explanation of this figure, mediation technology consists of a device for capturing gestural or motor-sensory information. As discussed this can be either haptic or non-haptic. The data is then emitted as gestural information. Depending on the device used, this can take many forms, but is generally obtained as some form of signal. The signal is then sent for processing into a useable form the system can utilise such as feature extraction in the case of images or variations in current, voltage or resistance over short distances such as those used in the bend sensors,

gyroscopes and accelerometers mentioned in sections 2.1.1.1 and 2.1.2.1 of Chapter 2. In terms of the technology developed for *In-Gest*, this method created the following system:

Figure 9. - *In-Gest* Mediation Model



Part of the information present in this model is obtained from the outline of the technology presented in section 3.1 of Chapter 2. The new element is *Procedural & Data-Drive Musical Events*, which is only possible where the mediation technology is applied in a heuristic sense to the compositional problem with the intention of exploring the gap between 'mind and matter' through the gestural affordances of the electronic score. This is how a mediation technology that acts as an 'action tracker' can assist in mediating cognitive information, especially in abstract concepts where the reception of ideas requires them to be processed before they are perceived. What is meant here is that the mediation form acts as an intermediary trigger of information (the linking of gestural affordances between two parallel musical scores) that is formed through cognitive processes. The trigger itself occurs through cognition related to

kinesthetic processes controlled by the motor-sensor system. This relates to the notion as depicted in section 1.1 of Chapter 2, where technology can act as both a mechanism of poiesis and esthesis. Furthermore, the development of a mediation technology to provide these problem-solving abilities in the context of composition, places the composer (author) in a position of a technologist where only minimal external expertise has been utilised.

CHAPTER 5 - A MEDIATION HEURISTIC TAKES FORM

*A composer who knows exactly what he wants, wants only what he knows—and
that is one way or another too little.*

– Helmut Lachenmann, 1999

5.1 - Development of the Mediation Technology for *In-Gest*

The mediation form proposed in the previous chapter is now explored where each aspect of the process as observed and experienced is detailed. The process revealed itself through research, experimentation and testing. At the outset, it was not known what the exact details of each section of the mediation form would be, nor how this would work within the process of musical composition itself. This was in-spite of the information gathered in the tool identification phase as detailed in section 3.1 of Chapter 2. However, with the application of requirements based on the content of the compositional problems as identified, the finer details of the structure became apparent. The Mediation Models (Figures 8., and 9.) as formerly proposed, refer to the macro structure of the mediation form and the sub-headings are taken from the model as per Figure 8.

5.1.1 - Devices and Capture Methods

In section 3.1.2 of Chapter 2 the type and manufactured brand of capturing device was detailed, along with the minimum requirements of a camera system to function when applied in the Viola & Jones (2001) method. The purpose of the section here is to describe some of the specifications and features and how they were applied in the mediation form.

According to the documentation available on the website⁴⁴, the Unibrain® Fire-i™ has a VGA 640x480 resolution with ¼" progressive image sensor. This allows for "individual readout of the image signals from all pixels" (Gasteiger, 2010:60). This seemed an important specification, as the successful application of responsive object tracking with the adopted method uses an algorithm that looks for the "sum of image pixels within specified rectangular areas" (OpenCV Development Team, 2014:448). As detailed in section 3.1.2 of Chapter 2, the subsequent camera acquired with higher specifications also utilises progressive scan image sensors.

5.1.2 - Gestural Information

This section of the mediation form is merely the information captured by the camera system before 'Processing'. This comes in the form of co-

⁴⁴ See <http://www.unibrain.com/products/fire-i-digital-camera/> for details (accessed 2014, November 15).

ordinate data that is a result of a complex process involving several different software applications and methods (detailed in 1.3 below). The information is naturally related to specific elements in the musical score where as discussed, musical gesture creates instrumental gesture. This relationship has been placed within the overall gesture space and information theory as in section 1.2 of Chapter 4 with the *Music Mediation Model* as proposed.

5.1.3 - Processing

There are two software applications utilised in the mediation form as intimated in section 3.1.1 of Chapter 2. They are discussed below, where their application, structure, details and methods are outlined. It should be noted that all the PD patches were created from scratch by the author, unless where indicated such as abstractions or add-on patches used that are available in the PD online repository⁴⁵.

5.1.3.1 - Pure Data (PD)

The first software application considered is PD. Section 2 of this chapter deals with a break-down and outline of the patches as applied in the work *In-Gest*. For the purposes of the discussions that follow, a brief terminology is necessary contingent with the nature of software

⁴⁵ See <http://forum.pdpatchrepo.info/category/2/patch> (accessed, December 13, 2014).

programming when using PD - see Table 2., below.

Table 2. – PD Terminology

Term	Description	Function
Objects	Displayed as a box in the GUI of PD and in text as [object], where <i>object</i> is the name that describes function.	A virtual processing unit that enables input and output. Processes signal or data rates where signal (audio) is prefaced by ~. There are two types of objects namely externals and internals. Externals provide functionality developed as add-ons whereas internals are necessary parts of the core PD program. ⁴⁶
Messages	Displayed graphically as a box with an open bracket right-side edge. In-text descriptions are presented as [message(, where <i>message</i> is either numerical (float) or a symbol (text).	They are used to input and store [message(data. Messages are essentially events that are sent between [objects] and can be floating point numbers, lists, symbols etc. ⁴⁷
Patch	Symbolically, <i>patching</i> refers to the connections made between pieces of equipment in a music	All the information created in the GUI window to perform the functions are stored in a text-based file

⁴⁶ See Farnell (2010:152).

⁴⁷ See Farnell (2010:158).

Term	Description	Function
	<p>studio. The PD program is designed to be representational of patching studio gear.⁴⁸ Besides the obvious reference here, <i>patch</i> is also the term used to describe the files created in PD.</p>	<p>given the extension .pd. These are referred to as <i>canvases</i> initially. Patches created in same instance of PD (additional canvases) can be linked through internal objects such as [send] and [receive].⁴⁹</p>
Sub-Patch	<p>An object created in PD that creates a new canvass in the form of a patch. In the GUI it takes the form of [pd xyz] where <i>xyz</i> is the name of the sub-patch.</p>	<p>The purposes of sub-patches are to limit the clutter of the parent patch where complex processes require multiple canvases. Communication takes place through [inlet] and [outlet] objects that are specified as either control or signal rate.</p>

Processing in PD takes place in two simultaneous parts or layers. They are the signal and control components as discussed in the table. This takes the form of a scheduler where DSP objects are interpreted (Farnell, 2010:150). Dataflow is achieved by communication along virtual lines displayed in the GUI as connections.

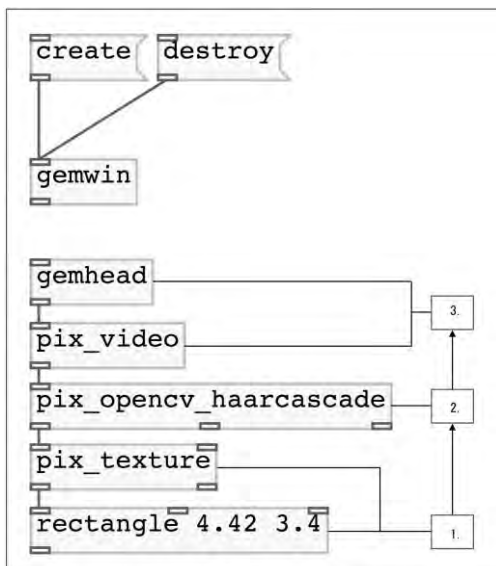
⁴⁸ See Chung (2013:7).

⁴⁹ See Farnell (2010:157).

5.1.3.1.1 - Graphics Environment for Multimedia (GEM)

GEM is an external for PD that provides the program with the possibility to convert data into image matrices. The external provides support for live camera feed via the [pix_video] object. This is achieved by objects added to a 'render-chain' to allow for adaptability and features. Consequently, this gives GEM the ability to process multiple functions prior to compiling the image matrix. The possibility to use the OpenCV bindings developed by Gómez & Degoyon (2009) for object tracking in PD/GEM are now clear. Figure 10., below, shows the render chain pathway with the OpenCV external object in the chain:

Figure 10. - GEM Render-Chain



In PD the processes are considered on a depth-first basis. It is important to be aware that the video stream occurs the instant the [pix_video]

object is created on the canvas without it being connected to the [gemhead]. We only see it once it is connected and rendered. The [gemwin] object creates a window where the image matrices are sent and is a separate process from [gemhead]. Following the objects from the bottom up, 1., a rectangle of (x,y) dimensions is created on a texture using [pix_texture] which is then 2., 'mixed' with the data stream from [pix_opencv_haarcascade] ⁵⁰. This object performs the tracking and recognition of the item in question (finger-tips), by using the trained classifier file created in training - see section 1.3.2.1 below. This information is 3., simultaneously applied to the video stream via [pix_video]. The [gemhead] object allows for an image layer to be textured on the [gemwin] canvas, which contains all the data. For a detailed reading of the workflow and programming idiosyncrasies of PD, the reader is directed to Puckette (1997), Kreidler (2013), Farnell (2010) and Chung (2013).

5.1.3.2 - OpenCV

OpenCV is the second software application in the system. Section 3.1.1 of Chapter 2 dealt with an introduction and contextualisation of this application in the heuristic. The details that follow intend to shed light on the workings as applied.

⁵⁰ This PD object was developed as a binding for OpenCV and PD by Gómez & Degoyon (2009).

5.1.3.2.1 - Haarcascade Classifiers

The term *haar* is applied to the process where the method adopted by Viola & Jones (2001) resembles haar-wavelets, a system of signal decomposition. The reader is referred to Papageorgiou et al., (1998) for further information. A *classifier* is a cascade of boosted classifiers (information that is concatenated in stages) that work with haar-like features and trained with sample images (OpenCV Development Team, 2014:447). The structure of the algorithm and its specific features compare images specified by shape and position within the ROI⁵¹. For a complete and thorough overview of the algorithm, the reader is directed to Viola & Jones (2001:514), the OpenCV Reference Manual⁵² and Dawson-Howe (2014:157).

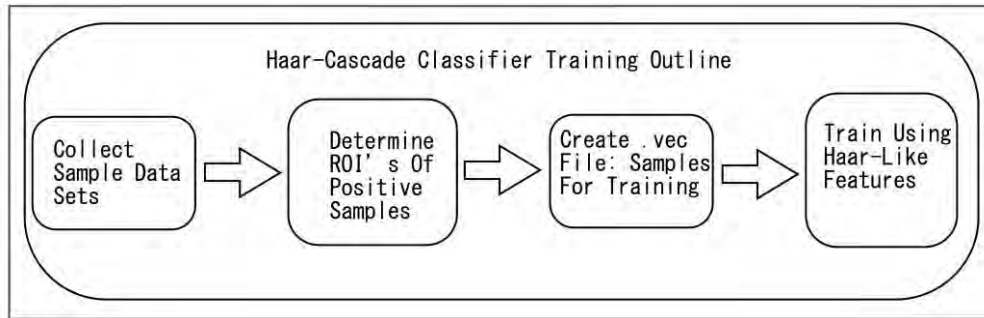
5.1.3.2.1.1 - Training Method Outline

An outline of the training process and method is given here to contextualise the process prior to the application of it for the creation of training data necessary for the heuristic as applied in *In-Gest*. The entire process took place over the period of just under a year between 30th September 2013 and 28th August 2014. Figure 11., below, summarises the steps:

⁵¹ See section 3.1.1 of Chapter 2.

⁵² Refer to <http://docs.opencv.org/opencv2refman.pdf> (accessed 2014, June 2) for details.

Figure 11. - Haar-Cascade Classifier Training Outline



Each of these stages are now detailed below:

5.1.3.2.1.1.A - Collect Sample Data Sets

The sample collection process involves gathering many thousands of images. They are collected in three sets namely positive (pos) and negative (neg) with a final set created for training purposes.

Positive Samples

The positive image set must contain the objects needed for detection. These are later added to the negative samples in the form of a training positive sample-set with distortions applied for training purposes (Seo, 2007). In the case of the mediation technology used here, these are finger-tips. This was achieved by taking a 5 second video of different left-hands supplied by willing contributors⁵³. Each hand was placed on a black background and filmed using an iPhone 5 resting on top of a camera

⁵³ Refer to Appendix IV for the list of contributors.

stand. The frames were then extracted from each video clip using an application called MovieDek⁵⁴. This generated 150 images per hand, and total of 13 different left hands were used in the pool giving 1950 images of hands. The next task necessary was to isolate the ROI, which in each hand are the finger-tips. Since the data-collection method described by Seo (Ibid) was adopted, it became apparent that the best detections by the system are achieved by positive images where the ROI is the size of the extracted image. By selecting a single-hand from each of the 13 pools to be used as a source for finger-tips, 4 fingers from left to right in each image were cut using a bulk image cropping tool called CropIt!⁵⁵. This also ensured that each image was identical in size - a requirement of the system as per the OpenCV Development Team (2014:447) - and therefore the ROI co-ordinates for each image would be identical (40 X 40). The overall process here created a set of 7800 positive images with 4 different finger-tips per hand. Initially, it was thought that this would be a suitable data-set based on its size, a larger set increases quality of training and thus detection (Dawson-Howe, 2014:152), but subsequent testing and experimentation revealed that such a large positive sample set caused training stage sessions to last for many weeks at a time, with the application needing to be terminated prematurely. In the end, a positive data-set of 1760 finger-tip images was used to create a successful classifier. Refer to Table 3. - *Training Attempt Historiography*, for details.

⁵⁴ MovieDek, by Limit Point Software, is available from the Mac App Store for Mac OS X.

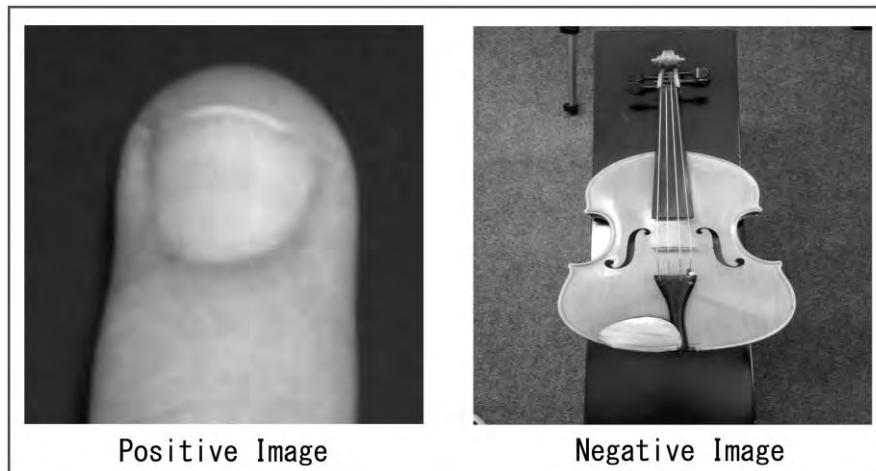
⁵⁵ CropIt!, by Totte Alm - Code 66, is available from the Mac App Store for Mac OS X.

Negative Samples

There are conflicting views as to the number of negative samples required for haar-classifier training purposes. This is because the number required is not listed in any of the available documentation. This seems to be apparent due to the open-source nature of OpenCV. With investigation, it was revealed that the numbers are determined in a comparison ratio with the quantity of positive images. Seo (2007) used just over 50% of the total number positive samples for the creation of their negative sample set. Whereas various discussions in the forums⁵⁶ revealed that the number should be at least double the total number in the positive set. It was decided based on the personal advice given to use a larger data-set. This particular process involved a painstaking, tedious and lengthy search for random images online, as well as filming scenes and extracting the individual frames that do not contain the object for detection i.e. finger-tips. There are some databases of images available such as those listed by Seo (Ibid). The negative data set images sizes may vary and should be converted to greyscale (Ibid). A total number of 4004 images were collected for the negative data-set which is just over double the number in the positive image set. Figure 12., overleaf, gives examples of the images from both sets.

⁵⁶ Refer to <http://answers.opencv.org/question/29126/haarcascade-finger-tip-detection-and-image-prep/?comment=29127#comment-29127> This discussion took place between the author and a member of the OpenCV online community in February 2014 (Warrington & Annamraju, 2014).

Figure 12. - Positive and Negative Image Examples



5.1.3.2.1.1.B - Determining ROI's of Positive Samples and vec File Creation

In these stages, the positive samples are applied to the negative samples with distortions added and then the ROI's determined. This is a critical stage where in a simplified explanation the system uses the training samples created to determine what is positive and what is not. It stands to reason the more negative images consumed, the better the detection. There are two utilities that provide functionality for the creation of the final training samples. They are `opencv_createsamples` and a C++ utility `mergevec.cpp` developed by Seo (2007). The `opencv_createsamples` utility is applied in two different ways here shown below as an extract from the Terminal during processing. They are delineated by A and C respectively, where A is the creation of individual image files and C, the final training samples file are created. B represents the merging of the

individual files. These take the form of the following terminal command lines as entered:

```
(A) opencv_createsamples -bgcolor 0 -bgthresh 0 -maxxangle 1.1 -maxyangle  
1.1 -maxzangle 0.5 -maxidev 40 -w 40 -h 40 -img pos1277.JPG -bg tmp -vec  
samplesos1277.JPG.vec -num 5
```

Info file name: (NULL)

Img file name: pos1277.JPG

Vec file name: samplesos1277.JPG.vec

BG file name: tmp

Num: 5

BG color: 0

BG threshold: 0

Invert: FALSE

Max intensity deviation: 40

Max x angle: 1.1

Max y angle: 1.1

Max z angle: 0.5

Show samples: FALSE

Width: 40

Height: 40

Create training samples from single image applying distortions...

Done

```
(B) 68-5b-35-b3-2b-1c:haartraining mileswarrington$ perl  
createtrainsamples.pl positives.txt negatives.txt samples 4000  
opencv_createsamples -bgcolor 0 -bgthresh 0 -maxxangle 1.1  
-maxyangle 1.1 maxzangle 0.5 -maxidev 40 -w 40 -h 40  
68-5b-35-b3-2b-1c:haartraining mileswarrington$
```

```
(C) opencv_createsamples -info positive.txt -vec samples_40x40.vec -w 40 -  
h 40 -num 41676
```

Info file name: positive.txt

Img file name: (NULL)

Vec file name: vecfile.vec

BG file name: (NULL)

Num: 41676

BG color: 0

BG threshold: 80

Invert: FALSE

Max intensity deviation: 40

Max x angle: 1.1

Max y angle: 1.1

Max z angle: 0.5

Show samples: FALSE

Width: 40

Height: 40

Create training samples from images collection...

positive.txt(901) : parse errorDone. Created 3603 samples

Firstly, the `opencv_createsamples` utility adds the positive samples collected to the negative samples with distortions applied by creating multiple `vec` files. `Vec` files are “universal OpenCV structure(s) that are able to store a single instance of primitive data type(s)” (OpenCV Development Team, 2014:7), such as floats and unsigned characters. The parameter `-num 41676` is obtained by applying the following formula:

$$\text{vec} \geq [\text{numPose} + (\text{numStages}-1)] * [(1 - \text{minHitRate})] * (\text{numPose}) + S,$$

where S is a count of samples from `vec` file that can be recognized (sic) as background right away (Warrington & Annamraju, 2014).

It seems counter-intuitive that this data is required prior to the training stage, so it is obvious that these parameters should be considered well in advance and that the complete training process should be considered before creating the training samples.

During the creation of these files, a text file is created where the ROI of each of the positive samples is given. These take the following format as extracted from one of the examples created:

```
0001_0056_0349_0414_0414. jpg 1 56 349 414 414
0002_0214_0354_0568_0568. jpg 1 214 354 568 568
0003_0372_0156_0323_0323. jpg 1 372 156 323 323 etc...
```

The numbers are co-ordinates in the format starting with the number of objects then x, y, width and height of object in specified size 40x40 as above. The text file is used by OpenCV in the creation of the samples by using that data to create the final packed vec file.

This produced many thousands of training samples with the positive images applied to the negatives with distortions. Figure 13., below, shows some examples of these images created.

Figure 13. - Training Sample Examples



The multiple vec files are then packed together using the mergevec.cpp utility to create the final vec file needed to begin training. It should be noted that at this stage advice and assistance was needed to complete a successful creation of the final vec file for training. Many failed attempts to create a suitable final packed vec file were the principal

reason. The author elicited the expertise of a technologist in the robotic and electrical engineering fields with past experience in merging vec files. This was sourced online via the OpenCV Questions Forum as discussed earlier where an exchange took place to uncover the correct number of negative images needed (Warrington & Annamraju, 2014). The reader is directed to the expert's website for further information (Annamraju, 2014). The contribution is noted as applied where necessary.

5.1.3.2.1.1.C - Train Using Haar-Like Features

Training cascade classifiers in OpenCV requires the vec file created earlier and the original negative sample set without the positives applied. Firstly, a description file of the contents of the negative image set is created. This gives the locations and file names of the negative images to be consumed. The output format of the file created is shown below:

```
./neg/neg0001.JPG  
./neg/neg0002.JPG  
./neg/neg0003.JPG  
./neg/neg0004.JPG etc...
```

Training is undertaken where variables are assigned to commands entered in the terminal. The commands were entered in the following form to create the last successful classifier:

```
opencv_haartraining -data haar -vec samples_40x40.vec -bg negatives.txt -
nstages 10 -nsplits 2 -minhitrate 0.9 -maxfalsealarm 0.1 -npos 1700 -nneg
3950 -w 40 -h 40 -nonsym -mem 3072 -mode ALL
```

Where:

- `opencv_haartraining` is the executable used to start classifier training in OpenCV;
- `-data haar` is the output file in xml format used in PD by `[pix_opencv_harcascade]`;
- `-vec samples_40x40.vec` is the name of the final packed vec file to be used;
- `-bg negatives.txt` is the description file of the original unaltered negative images;
- `-nstages 10` is the number of stages set to train - higher number usually yields better results, but training may terminate if desired `maxfalsealarm` or `minhitrate` are achieved;
- `-nsplits 2` are the number of splits in the tree of cascades - according to Seo (2007), `nsplits < 3` works best for samples 20 X 20 and over;
- `-minhitrate 0.9` where a minimum hit-rate of 0.9 is set. This parameter sets the minimal desired hit rate for each stage of the

classifier;

- `-maxfalsealarm 0.1` in essence this sets the percentage (1%) of allowed false positives;
- `-npos 1700 -nneg 3950` where each of these are the specified number of positive samples used in the packed vec file and number of negative images to be consumed from the original negative samples data-set;
- `-w 40 -h 40` - specifies the width and height of the sample size used in the creation of the vec file - they must match;
- `-nonsym` is an important parameter since it specifies the training of non-symmetrical objects where there is no left-right (vertical) symmetry. Finger-tips are obviously asymmetrical;
- `-mem 3072` specifies the maximum amount of system RAM should be allocated during training - in this case 3072 MB;
- `-mode ALL` sets the training to consider extended sets of Haar-like features and will incorporate the distortions created in the `opencv_createsamples` utility.

The parameters for training are gleaned from the number of positive and negative images used to create the training samples. Standard training practice allows for 90% of the positive samples used in the `opencv_createsamples` utility hence `-npos 1700`. Similarly, slightly less than the total number of negative images used in the utility are applied

in training, hence -nneg 3950. The reason for this is to allow some spare images for consumption.

The training of this classifier was the 25th attempt (see Table 3., on page 114), and took place over approximately 4 weeks in August of 2014. A sample of a stage is included in Figure 14., overleaf. The training terminated after 5 stages and yielded an xml file called haar.xml, which was used in conjunction with PD. The xml is included in Appendix III in data format on the DVD ROM supplied.

The table that follows Figure 14., details the complete history of training attempts before a usable classifier was created where an accuracy of approximately 70-80% depending on lighting conditions was achieved.

Figure 14. - Stage in Final Trained Classifier

```

Terminal Saved Output

*** 1 cluster ***
POS: 1700 2350 0.723404
NEG: 3950 3.02661e-05
BACKGROUND PROCESSING TIME: 2043.00
Precalculation time: 1.00
+-----+
| N |%SMP|F| ST.THR | HR | FA | EXP. ERR|
+-----+
| 1|100%|-|-0.455708| 1.000000| 1.000000| 0.175929|
+-----+
| 2|100%|-| 0.019156| 0.946471| 0.429367| 0.316283|
+-----+
| 3|100%|-|-0.445110| 0.970000| 0.463038| 0.120531|
+-----+
| 4| 79%|-|-0.033627| 0.944706| 0.239747| 0.123894|
+-----+
| 5| 82%|-| 0.311582| 0.902353| 0.051899| 0.086372|
+-----+
Stage training time: 7929.00
Number of used features: 10

Parent node: 3
Chosen number of splits: 0

Total number of splits: 0

Tree Classifier
Stage
+-----+
| 0| 1| 2| 3| 4|
+-----+

    0---1---2---3---4

Parent node: 4

```

Table 3. - Training Attempt Historiography

Attempt	Date	Vec Size	NumStages	Outcome : NumStages	Usable xml?
1	30/09/2013	30X20	5	Errors in createsamples	x
2	01/10/2013	190X260	5	Errors in createsamples	x
3	11/11/2013	70X171	10	Auto-Terminate: 2	x
4	11/11/2013	20X60	10	Auto-Terminate: 2	x
5	12/11/2013	54X96	10	Auto-Terminate: 2	x
6	18/11/2013	108X192	10	Auto-Terminate: 1	x
7	21/11/2013	29X84	20	Error in dataset	x
8	23/11/2013	29X84	20	Manual-Terminate: 2	x
9	27/11/2013	12X34	20	Manual-Terminate: 9	x
10	07/01/2014	12X34	20	Manual-Terminate: 8	x
11	07/01/2014	12X34	10	Auto-Terminate: 6	x
12	09/01/2014	13X24	10	Auto-Terminate: 8	x
13	10/01/2014	12X34	10	Auto-Terminate: 9	✓
14	25/02/2014	27X48	10	Auto-Terminate: 4	x
15	24/03/2014	60X80	10	Manual-Terminate: 3	x
16	12/05/2014	24X24	15	Manual-Terminate: 4	x
17	14/05/2014	27X48	5	Auto-Terminate: 4	x
18	18/05/2014	27X48	10	Manual-Terminate: 4	x

Attempt	Date	Vec Size	NumStages	Outcome : NumStages	Usable xml?
19	27/05/2014	20X20	10	Manual-Terminate: 3	x
20	30/05/2014	20X20	15	Auto-Terminate: 9	x
21	30/05/2014	20X20	15	Auto-Terminate: 2	x
22	31/05/2014	20X20	15	Auto-Terminate: 3	x
23	06/06/2014	20X20	15	Auto-Terminate: 4	x
24	10/06/2014	20X20	15	Manual-Terminate: 4	x
25	01/08/2014	40X40	5	Auto-Terminate: 5	✓
26	28/08/2014	160X160	5	Manual-Terminate: 4	x

5.1.4 - Summary

It is apparent from this historiography that patience is certainly required in training classifiers. Every attempt utilised differing parameters, including image sizes and sets, where -nstages, -nsplits, -minhitrate, -maxfalsealarm, -npos, -neg etc., each received alternative parameters before a successful xml was created in attempt 25. The 40X40 vec at just over 160MB in size yielded an approximate accuracy of 70-80% successful detection rates of the objects dependent on lighting conditions and angle of camera to object plane. It was discovered that direct fairly bright light illuminating the fingers with the camera positioned parallel and up to a maximum of 45 degrees to the fingers is optimal. The camera's field-of-view should also be positioned such that the finger-tips are not

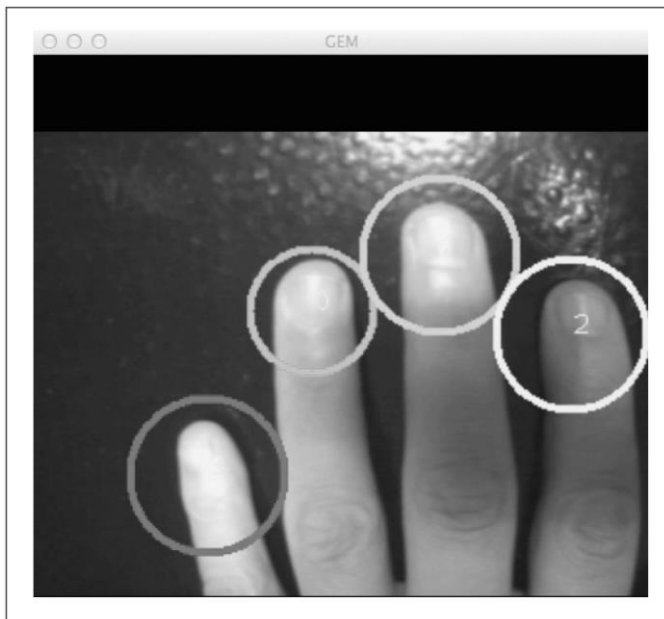
upside-down in the rendered image. It was discovered that having them upside-down depleted detection accuracy formidably. Furthermore, the inability for the system to make use of the higher specification camera as detailed in section 3.1.2 of Chapter 2, means that the detection accuracy and rate is further reduced. This had an overall effect on the facility for [pix_opencv_haarcascade] to prevent multiple detection instances of the same object being given differing identifiers. In other words, every-time a detection of the same finger-tip was dropped due to inaccuracy, the system would give the finger-tip a different identifier. This proved to be a major drawback of the system in its current state, as it has direct implications of the ability for the *In-Gest* Audio Patch to process the detection points accurately. The shortcomings are given attention in section 2.1.4 of this chapter and again in summation in section 1.4 of Chapter 7. Figure 15., below, depicts the most successful setup and position used for demonstration purposes.

Figure 15. - Optimal Camera Setup



The accuracy of the 160X160 vec file would have been significantly higher in attempt 26, but this was manually terminated after the 6th week of constant training with no result forthcoming and the computer overheating. The vec file for attempt 26 is over 500MB and requires a multi-core computer array in order to be processed more efficiently and effectively. The process to create this final vec file was assisted with information and partnered training attempts by Warrington & Annamraju (2014). The successful training of a classifier file using the larger vec file is certainly a task worthy of attention in continuation and future developments of the heuristic. Further possible improvements with the application of the higher specification camera and re-designed OpenCV binary externals for PD that limit object identifier changes once detected, would certainly yield more stable object detection and better performance results. As is documented in section 2 of this chapter, the shortcomings of the current system require that a controller is present during a performance or rehearsal of the work to manually perform audio processing control should the detection and tracking not execute correctly. In fact, this was essential in the demo recording session. Figure 16., overleaf, shows a full hand result of the detection taking place in the PD/GEM video patch applied.

Figure 16. – Full Hand Finger-Tip Detection in PD/GEM



5.2 - Application of the Mediation Technology in *In-Gest*

As discussed in section 2.2 of Chapter 4, the application of the system in the musical work *In-Gest* required heuristics whereby the objects tracked would provide scope and possibility for the triggering and manipulation of audio content in live performance. This process is detailed below in a pictographic form where parts of the PD patches are detailed and explained.

The intention of this section together with the video clips of a live studio demo recording in Appendix III is to provide a demonstration of proof-of-concept for the mediation form as applied in the compositional process, resulting from the musical ideation as discussed. In essence,

this is a visually documented reference of the unfolding of the composer-technologist occurrence. Chapter 6 deals with compositional elements pertaining to the music.

5.2.1 - *In-Gest* PD Patches

In order to comply with the necessity to avoid audio jitter or stutter during simultaneous audio and video processing in one session of PD (refer to previous discussion in section 3.1.1 of Chapter 2), two separate patches were created.

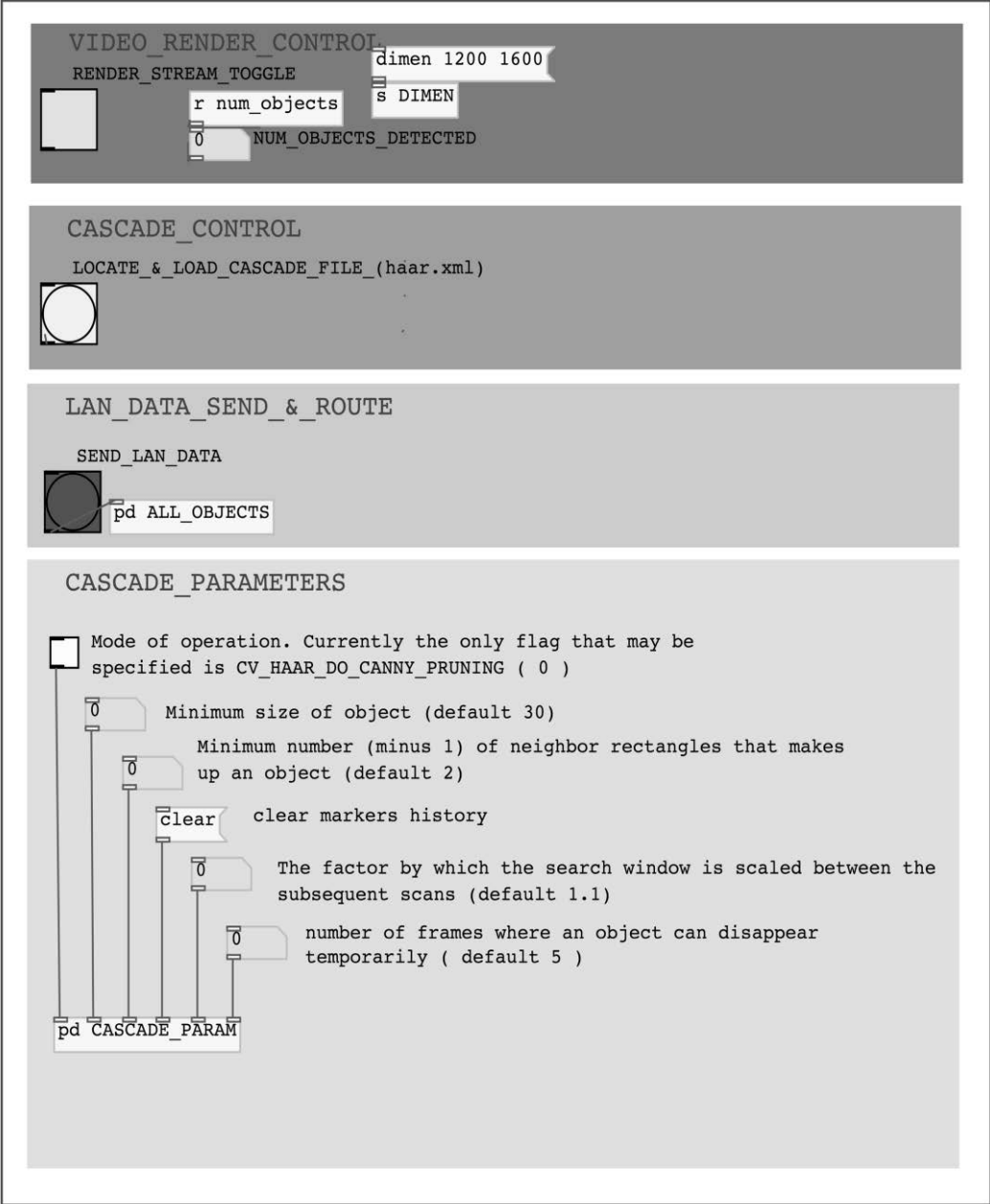
5.2.1.1 - *In-Gest* Video Patch

The overall structure of the video processing patch is outlined below, where each figure is an exported post-script file from PD where the main patch and all its sub-patches are shown. Each figure is provided with an explanation pertaining to the function and application of the patch. Figure 17., on the next page, is an export image of the main video patch, which is divided into several 'control' sections.

VIDEO_RENDER_CONTROL at the top of the patch initialises the [gemwin] object and outputs object detection data through a sub-patch - see sub-patch section below. CASCADE_CONTROL allows the controller to locate and load the trained classifier xml file (haar.xml). LAN_DATA_SEND_ROUTE

section contains sub-patches that route the object tracking data (x, y coordinates; z size; object identifier number) to the computer hosting the audio processing patch via a LAN connection. CASCADE_PARAMETERS allows the user to apply various settings to the [pix_opencv_harcascade] object that performs the object detection with the use of the haar.xml file.

Figure 17. - *In-Gest* Main Video Patch

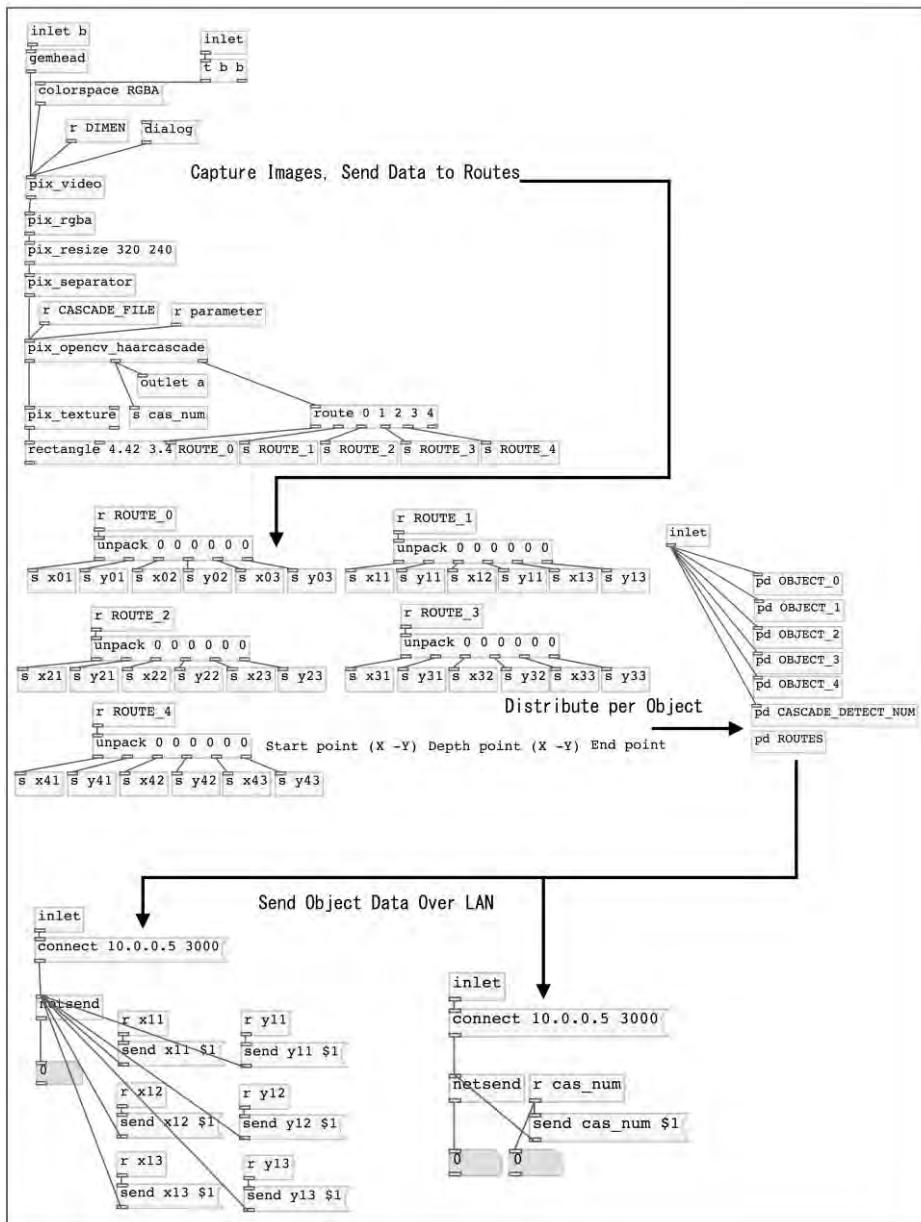


5.2.1.1.1 - *In-Gest* Video Sub-Patches

Most of the sub-patches present in the main video patch perform the primary function of routing data from the video stream. Each object detected outputs three parameters: start-point, depth-point and end-point. Furthermore, each of these contain (x;y) co-ordinates.

Figure 18., overleaf, shows all the sub-patches with their dataflow in perspective and relation to each other. This provides a top-down-view of the flow of the data from finger-tip to LAN connection. Each of the possible 5 detectable objects (thumb included) is given a set of variables: (X01; Y01; X02; Y02; X03; Y03). The 1st numeral corresponds to the object number (identifier from [pix_opencv_haarcascade]), with the 2nd numeral indicating their role such that 1 = start-point, 2 = depth-point and 3 = end-point. X and Y are their relative co-ordinates. The numbering starts at 0, since 0 is considered an identifier by [pix_opencv_haarcascade]. This means there are 0-4 (inclusive) possible data routes, each allowing the allocation of specific sets of variables per identified object to be routed. The main audio patch is given an outline next and is detailed more thoroughly in Chapter 6, where each detection point is discussed in the musical score.

Figure 18. - Video Sub-Patch Network



5.2.1.2 - *In-Gest* Audio Patch

As discussed previously, the audio patch was run on a separate machine being fed the capture data from the station running the video patch over a LAN connection. The diagram in Figure 20., on page 126, shows the setup used in the demo recording session, which was used to capture live

detection in a few examples for a proof-of-concept demonstration (See Appendix III - DVD DISC). The audio processing patch performed two macro functions:

1. stream audio data at specific trigger points in the score off the hard-disc drive (data-driven events);
2. perform live manipulations and transformations where required in the score (procedural events).

The printed violin score is made-up of several single sheets or 'movements' that are playable in any order - see Chapter 6 for further details. Therefore, there are 7 sub-patches relating to the separate gesture sheets. The controller must enable each sub-patch during performance where the order is chosen prior to performance.

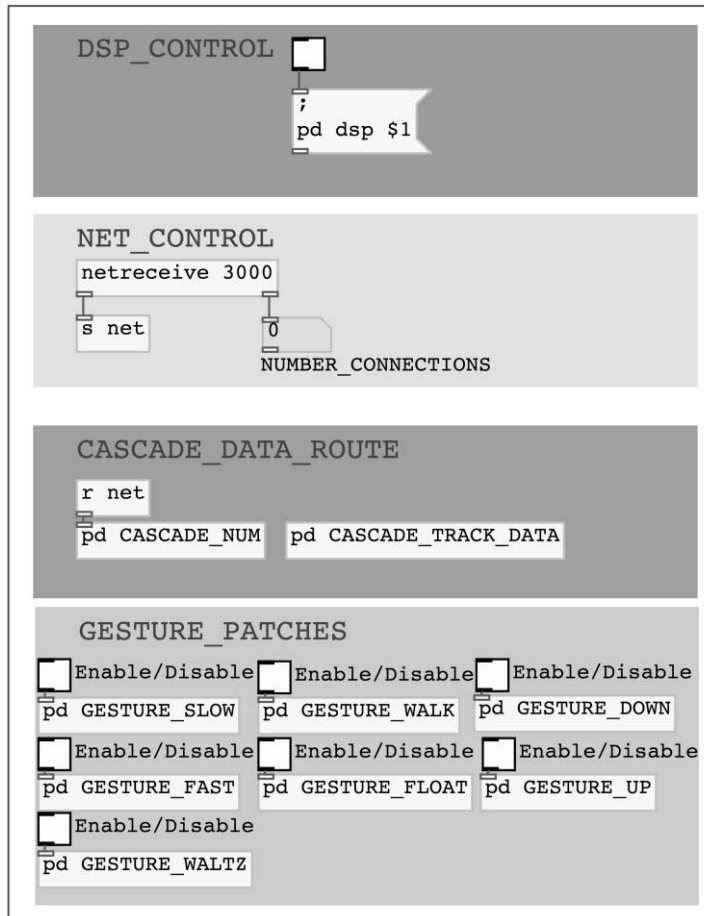
The nature of the data-stream from the video processing station forces this situation because multiple similar variables are captured for each patch that would be impossible to isolate for specific audio events if the patches were not enabled and disabled as needed, since a constant data-stream is occurring from the video processing station whilst the [gemwin] is rendered. It became apparent that due to the wide range of tracking data received per object and fed into the [pd OBJECTS_RECEIVE] sub-patch, to correlate tracking data to detection point would require the data to be contained within specific ranges for each detection occurrence. Achieving this necessitated separate control of the data-rate, before being

allocated to specific tasks via tailored PD [send XYZ] objects per patch (see section 2.3.2 of Chapter 6 for details).

Owing to the current inherent inaccuracy of the system, a direct audio control patch (panic patch) was created to allow the controller to perform manual audio streaming should the system miss a detection completely. Consequently, it is necessary for the controller to have a copy of the complete score.

Figure 19., on the next page, shows a graphic export of the main audio patch as applied. The main patch is divided into 4 main sections: DSP_CONTROL, NET_CONTROL, CASCADE_DATA_ROUTE and GESTURE_PATCHES. DSP_CONTROL initialises the audio render (digital signal processing engine) of PD. NET_CONTROL is a sub-patch that collects the data streaming from the Video Processing Station via LAN connection. CASCADE_DATA_ROUTE provides the critical distribution of tracking data to the GESTURE_PATCHES, which contain the parameters for both procedural and data-driven audio events as per the score. Each of the sub-patches that relate to the score where musical materials form a part of the patch's task, are detailed with excerpts from the score in Chapter 6.

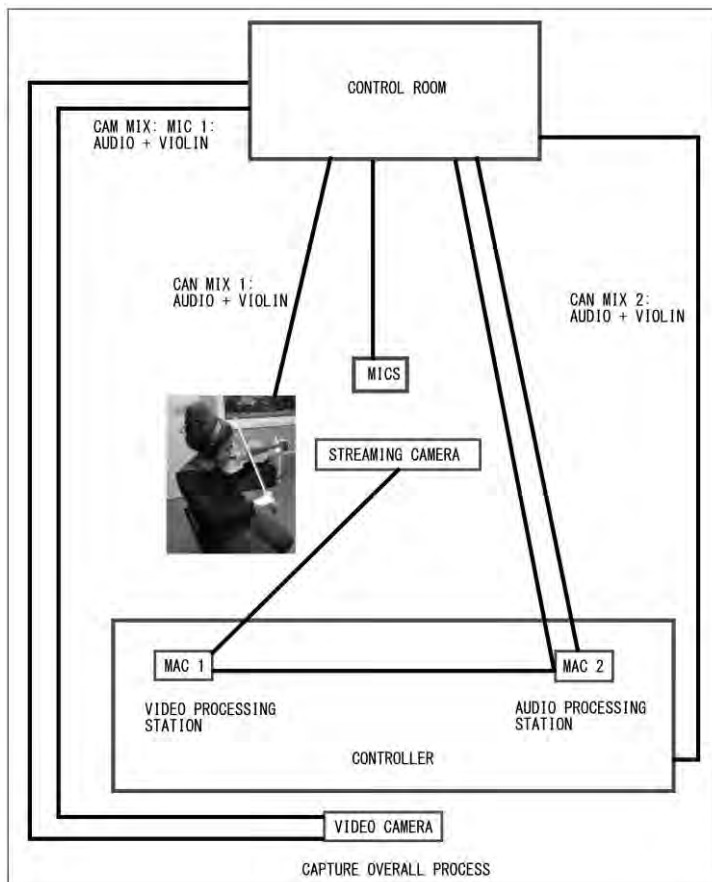
Figure 19. - *In-Gest* Main Audio Patch



5.2.1.3 - Panic Patch

As previously discussed, it was wholly necessary to provide the controller with a 'panic patch' whereby an alternative set of audio content (where procedural audio would be indeterminate it is now fixed), could be played back manually should the system fail to detect correctly or indeed miss a detection all-together. The patch sets triggers for the audio content without any procedural manipulation. See Figure 21., on page 128.

Figure 20. – Demo Recording Audio Setup and Layout



5.2.1.4 - Future Enhancements

It is hoped that as discussed in section 3.1.2 of Chapter 2, that the developers at IEM Graz do indeed fix the bug that causes the higher specification and natively supported fire-wire camera to crash PD.

Together with that, it is intended that a more stable and focused PD object that deals with the interpretation and live processing of the trained haarcascade classifier file ([pix_opencv_haarcascade]), be created to allow only single detections of any one object once detected.

As discussed, in its current form, the [pix_opencv_haarcascade] PD object identifies an object, gives it an identifier (number) but the connection is dropped at subsequent detections and it therefore has a tendency to lose this number. With the intended future enhancement that would negate this tendency for dropping, each of the finger-tips could be given a fixed number, such that for example (1) could be permanently assigned to the index finger (finger 1 in violin fingering practice), and so on. The patch shown in Figure 22., on page 128, consists of two parts, with the left-hand side showing the connection from the video processing end and the right-hand side the distribution of the data at the audio processing side.

This would allow the routing of detected objects and therefore the finger-tip tracking data could be allocated per finger-tip, removing the need for a generalised routing system with variable ranges as deliberated. Each of the parent audio patches (see Appendix II), have been given the required data routing capabilities, which can be added once the OpenCV binding PD object is altered and functioning as expected. This would significantly improve the overall performance and accuracy of the system as developed here.

Figure 21. - Panic Patch Main Window

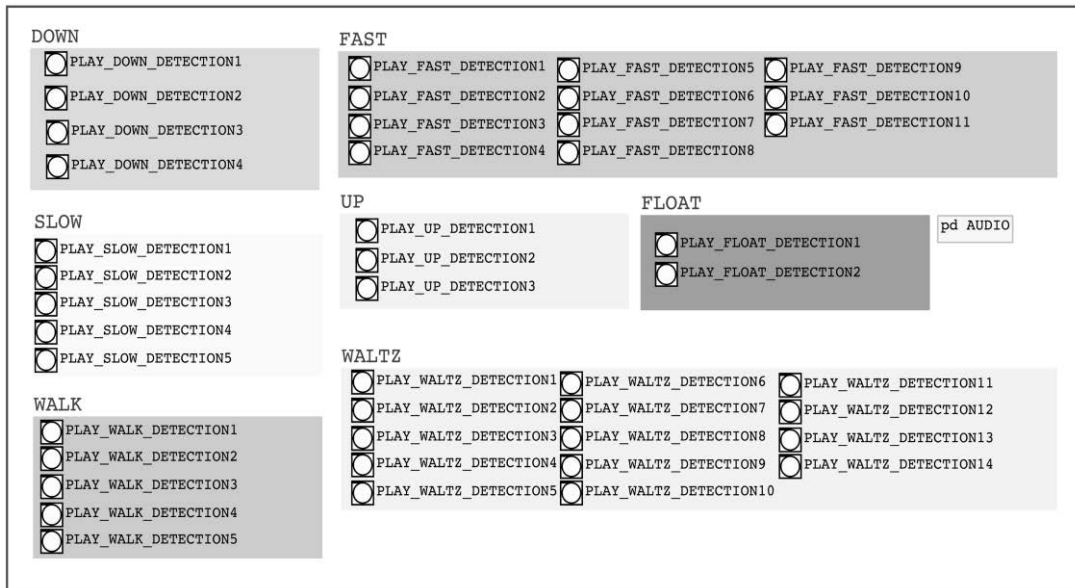
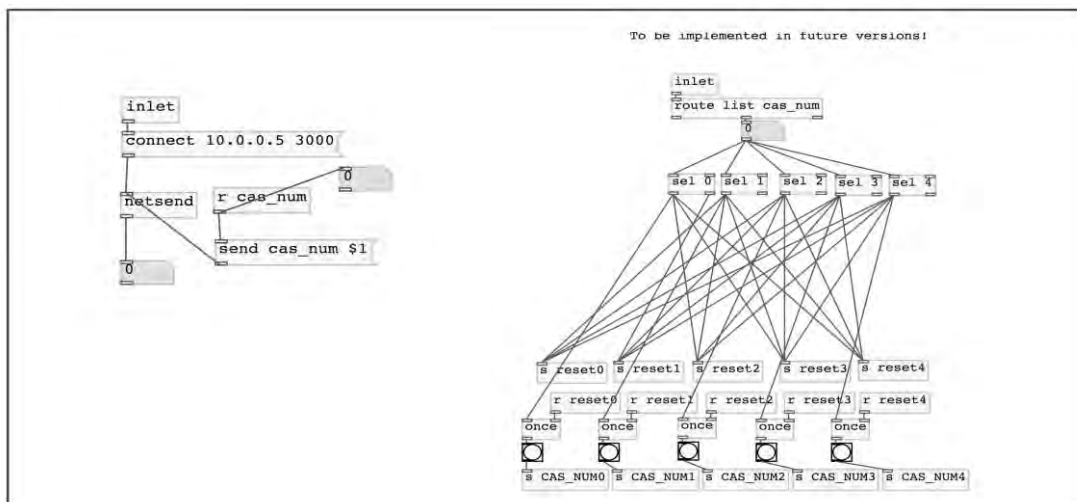


Figure 22. - Cascade Object Identifier Routes



CHAPTER 6 - THE COMPOSITION *IN-GEST*

Composing was always my major interest. I was always interested in a theory of composition...the process.

- *Otto Laske*

6.1 - *In-Gest*: An Introduction

Compositional process is by its nature idiosyncratic, with each experience and procedure of the individual not alike in the community of composers worldwide. The process for *In-Gest* has itself never been systematically linear, with the necessity to move between ideation, technique and implementation processes in a hermeneutic, multi-directional manner. Laske (1999:151) has commented that "dialog (sic) with oneself in art making is crucial, and that computer programs, rather than being simply tools for achieving results, are dialoguing tools." Naturally, some of these individual processes may indeed be more linear, such as the steps necessary to achieve functionality in part of the heuristic detail of the work. Or, the systematic approach applied to gestural signification such as in Models A and B as discussed. What is discussed here though, relates

to the creation of both the acoustic and electroacoustic scores for *In-Gest*, how the heuristic was applied to the scores and what the intentions for reception and aesthetics are in the work. This is undertaken within the overall process of the composer-technologist paradigm as considered and applied.

6.1.2 - Rationale Behind *In-Gest*

The primary objective of the research presented in this thesis is to detail a compositional process that applies a technologically centered gesture heuristic developed to that end to solve a compositional problem. *In-Gest* became the vehicle or laboratory for this process. Merely assuming the role of a 'test-tube' for process is not the only reason for adopting this method of experimentation, but rather to present it within the framework of a particular musical aesthetic and add to the result with a formed musical work. Perhaps significantly, 'rationale' as such is a questionable concept when applying it to composition because compositional intention, the creation of sound to illicit either internal or external emotional or intellectual responses, is rooted in a priori and is therefore justified independently of experience. The individual's perception in the act of reception is intentionally objective, however it may be 'guided' with techniques that create particular aesthetics. Essentially, the work attempts, as Hamman (2004:115) puts it in describing Xenakis' use of technology in music, to "project(ed) musical thought" in a

“technical experiment.” The notion of compositional ideation as previously discussed can now be placed within a practical framework, where in addition to the development of a gesture study schema and mediation models as proposed, the composer creates organised “musical *Gestalten*⁵⁷ in order to evoke impressions of motional–dynamic gestures, or even openly imitate such gestures through melodic movement, rhythmical patterns, etc.,” (Schneider, 2010:82).

6.2 - *In-Gest*: The Composition

As a work for violin and electronics, *In-Gest* can be analysed by considering both the acoustic and electronic scores simultaneously. The electronic score itself consists of two parts namely procedural and data-driven⁵⁸. ‘Procedural’ occurs where the timing of the elements in the electronic score are triggered by gestural events and indicated by specially tailored notational means. Where they occur as ‘data-driven’ elements, the score is ‘fixed’ and shown as such in the printed score. Whilst it would seem appropriate to split the analysis that follows in section 2.3.1 into these two categories, both the scores are not mutually exclusive since the work is conceptualised to afford musical gestures from one to the other. Furthermore, some elements in the acoustic score have a direct and visceral influence on the outcome of particular musical features where procedural elements occur, which are themselves pre-

⁵⁷ See section 1.1.2 of Chapter 4.

⁵⁸ See section 2.1.2 of Chapter 5.

composed but manipulated in real-time. For these reasons, the scores are detailed concurrently and in parallel. It is very important to consider the patches that perform procedural events as elements of the musical score, since they are generating musical material. The patches are 'composed' as they are constructed to perform a particular function musically, be it mimetic, gestural affordance, harmonic, rhythmic, relational or of other such important musical significance. Therefore, each section that contains procedural elements is discussed with further analysis of the patches created which work along with the heuristic 'behind the scenes'. The materials generated for the electroacoustic score are also given written attention.

6.2.1 - Composing Phrases for Implied Musical Gestures

'Implied' is given as a qualifier to the musical gesture definition here since a composer can only imply that the elements contain the gestures intended. To re-iterate, the musical gesture implied is an abstract compositional element or phrase that conveys a meaning. This type of musical gesture occurs in both the acoustic and electroacoustic scores, whilst they are mediated with the aid of an instrumental gesture heuristic in the form of an action tracker.

6.2.1.1 - Macro Structures: The Gestural Sheets

The gestural elements are formed in containers referred to as 'gesture sheets'. Each of these sheets have descriptive titles which imply a gesture. They are (in order of analysis): *Walk*, *Up*, *Down*, *Waltz*, *Float*, *Slow* and *Fast*. The titles could be thought of as gestural signifiers. Thus for example, *Waltz* and *Walk* each signify bodily movement in the form of a dance and gait respectively, whereas *Float* and *Slow* imply a sense of suspension but from differing perspectives. A similar interpretative framework is intended with *Up and Down*, where a steady sense of ascent and descent are implied. *Fast* on the other hand provides a sense of fleeting urgency. As described in Chapter 3, a further intention is to relate these movements to musical gesture, where the sound is "treated as image" (Mazzoli, 2014:24). Furthermore, the performer has the advantage of creating some form of narrative by choosing the order of the sheets. The performance length of each of the sheets could vary from reading to reading and is not intended to be fixed, however the work should not last more than 17 minutes in total.

6.2.1.1.1 - Pitch Classes

An important macro element is the use of a scale of 11 pitches:

Figure 23. - Pitches



Overall, these pitches form the bulk of the pitches used in the score, and no others are introduced unless they occur as part of procedural audio elements or where microtonality plays an important role in either of the scores. It is noticeable that it is almost a 12-tone chromatic scale, only one pitch (E) is missing. The choice to leave out the E was governed by the material collected and generated for the electroacoustic score. This process is outlined in section 2.2 of this chapter.

6.2.1.2 - Micro Structures

The micro structures contain the individual elements, themselves made-up of various compositional techniques and procedures to create specifically tailored sonic events. Each *gesture sheet* is considered with extracts from the complete score where important events occur, particularly specific detection points. The delineation is assembled in a diminishing format where larger elements are succeeded by smaller elements and so forth. There are no barlines in the work and the score is 'blocked' where no musical material occurs in a particular part. The lack of barlines is indicative of the intention that no formal grid or time structure is implied unless otherwise indicated. This was viewed as a compositional

advantage, a naturally occurring result of using a gesture heuristic of the nature applied, where timing between detection events forms structure rather than barlines.

6.2.2 - Generation of Electroacoustic Materials

The acquisition of materials in electroacoustic music is an important part of the overall compositional process and the techniques employed can have important and lasting effects on the structuring of a work. Sigal (2009:18) relates the acquisition of material to sound characteristics, where they influence the "language" of the piece. Indeed, much of the material generated for *In-Gest* resulted in important characteristics of the work, such as the omission of the pitch of E from the formal score. The materials were generated from several motives composed in the very early stages of the work and later served to generate larger structures for each of the gesture sheets. Additionally, several important extended techniques were recorded to allow for additional transformations of the techniques employed in the acoustic score. Importantly, the electroacoustic sound palette consists mostly of material recoded from pre-composed motives for the violin. This is in-line with a practice often employed in electroacoustic music creation such as demonstrated in the works of composer Denis Smalley, among many others. (Alcorn, 1992:47). Other sounds included are those of recorded footsteps on gravel, which are used in *Walk* with procedural transformations. The materials composed are

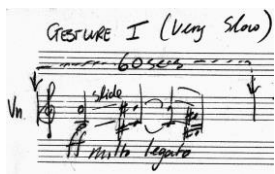
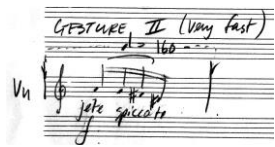
shown in Table 4. - *Elements for Electroacoustic Material Generation* below, where the original composed hand-written materials are included.

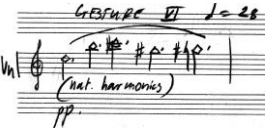
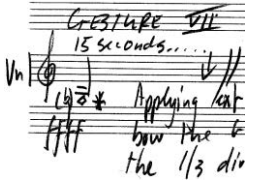
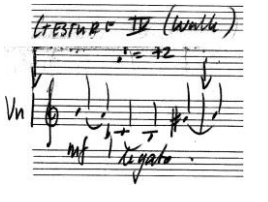
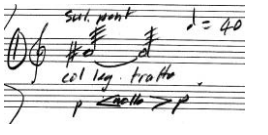
The material generated was recorded at the date indicated using the following equipment and setup:

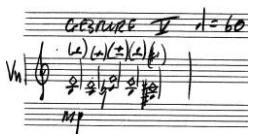
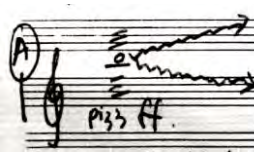
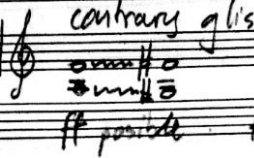
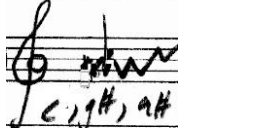
23 April 2014 (violin):

- 2x AKG C414 XLS in Spaced Pair configuration and Wide Cardioid pickup pattern
- 2x Rode NT55 in Coincident Pair configuration and Cardioid pickup pattern
- 1x Rode K2 with a Cardioid pickup pattern

Table 4. - Elements for Electroacoustic Material Generation

Original Composed Element	Transformations Applied	Gesture Sheet
	<p>Time-stretching in procedural events; Octave transpositions + & -; Harmonics & partial extraction/filtering; Layering.</p>	<p><i>Slow, Float</i></p>
	<p>Layering to form chord structures; Waveform direction reversal; Applied to procedural time-morphing events.</p>	<p><i>Fast, Waltz</i></p>

Original Composed Element	Transformations Applied	Gesture Sheet
 <p>Gesture VI 2/4</p> <p>nat. harmonics</p> <p>pp</p>	<p>Equalisation applied to extract harmonics;</p> <p>Octave transpositions + & -;</p> <p>Time-stretching in data-driven events;</p> <p>Layering.</p>	<p><i>Float, Up</i></p>
 <p>Gesture VII</p> <p>15 seconds...</p> <p>Applying cat</p> <p>bow the 1/3 dir</p> <p>fff</p>	<p>Sub-harmonic processing;</p> <p>Time-stretching in procedural & data-driven events;</p> <p>Multiple octave transpositions;</p> <p>Layering & distortion.</p>	<p><i>Down</i></p>
 <p>Gesture IV (Waltz)</p> <p>mf</p> <p>legato</p>	<p>Harmonics & partial extraction/filtering;</p> <p>Layering to create complex chord structures;</p> <p>Multiple octave transpositions & tuning;</p> <p>Filtering & equalisation.</p>	<p><i>Walk, Waltz</i></p>
 <p>Seri punk</p> <p>col leg. tratto</p> <p>p >>p</p>	<p>Layering & transpositions;</p> <p>Time-stretching.</p>	<p><i>Walk</i></p>

Original Composed Element	Transformations Applied	Gesture Sheet
	<p>Harmonic and partial extraction/filtering;</p> <p>Layering to create complex chord structures;</p> <p>Multiple octave transpositions;</p> <p>Procedural time-stretching.</p>	<p><i>Walk, Float</i></p>
	<p>Bi-directional pizz. layered;</p> <p>Time-stretching without effecting pitch applied;</p> <p>Multiple micro-tonal transpositions;</p>	<p><i>Float</i></p>
	<p>Layering;</p> <p>Multiple transpositions by semitones;</p> <p>Time-stretching in procedural events;</p>	<p><i>Walk</i></p>
	<p>Equalisation;</p> <p>Attack & Decay Envelope Transformations.</p>	<p><i>Waltz</i></p>

6.2.3 - Scores in Process

All the recorded material was carefully categorised and then assembled into useable forms by the methods of transformation as detailed. The characteristics of each of the forms provided important aesthetic guidelines in the process of musical assembly, where the relationship between the emerging acoustic score and electroacoustic materials

solicited imperative criteria in selection and control. Example criteria included such aspects as the nature of a time-stretched glissando; did they create descending harmonics that relate to the slides and glissandi present in the acoustic score, or have the transpositions provided adequate impetus to the intention of descent or ascent? In essence, it was imperative to ascertain how such relationships could be adequately applied in the score through the methods of notation and by 'scoring' with the creation of procedural patches in the heuristic. Such processes can be described by compositional techniques. The techniques used and applied in the creation of the scores for *In-Gest* include:

- mimesis
- overlaying
- assimilation
- relating differing sound materials
- inside versus outside time structures
- articulations between acoustic and procedurally created sounds
- overall gestural links - relationship between phrases that afford mental representations of sound between both scores

The somewhat unconventional nature of electroacoustic music where it requires scoring in the traditional sense runs into difficulties and challenges. Traditionally, electroacoustic works that have no live instrumental part have no notation, other than perhaps instructions for

playback and setting. Similarly, where works have instrumental involvement, it is often the case that the electroacoustic part does not have an accompanying notated score. As Risset (2002:xv) states, "Most music compositions for tape do not come with a score." The compositions of complex musical structures using computers or workstations have historically shown traditional scoring and notational methods to be archaic, perhaps even useless, such as in purely *musique concrète* works. Notwithstanding this, the lack of printed scores or even any analytical details at all present in most of the large opus of sophisticated computer generated music since composer and 'father of computer music' Max Matthews developed the first synthesised sound at the Bell Laboratories in 1957, is well documented (Risset, 2002:xiv). This has left a tremendous gap in the body of analytical literature undertaken by musicologists, because they rely on the objectification of process during the investigation of musical works (Ibid). However, with much work undertaken fairly recently in remedying the lack of electroacoustic music analysis, various techniques and methods have been applied with great success in producing meaningful objective representations of works that traditionally have no score. For a look into these methods as developed and applied, the reader is directed to the pioneering early work of Cogan (1984:103-119) and more recent analyses in Licata (2002), Simoni (2006) and Mazzoli (2014). Of interest is the forthcoming volume due soon and entitled *New Multimedia Tools for Electroacoustic Music Analysis* (Emmerson & Landy, 2012). On the notation front, Dimpker (2013) provides techniques, tools and methods for notating

electroacoustic music. Some of these techniques have been utilised in the creation of the scores for *In-Gest*, such as applying sonograms for sound analysis and interpretation. However in most instances, several new notation symbols have been developed to provide specific and particular indication of compositional intention in the score. Most of these use established means where minor manipulations or additions to notational conventions have been created and where done so, are detailed in the *Notation Guidelines and Symbols* section of the full score provided in Appendix I. The reader is referred there for its use in conjunction with the analysis that follows.

6.2.3.1 - *In-Gest*: An Analysis

The analysis undertaken here proceeds from a necessarily subjective perspective. This perhaps has both advantages and disadvantages. The advantages are clearly apparent in the development of a score for the electroacoustic part, which is an unusual but critical component of the work presented here. This negates the need to rely on undertaking analysis from a frequency versus time method such as the techniques developed by Cogan (1984) to create and analyse sonograms. Similarly, it removes the necessity to depend on analysis from a purely *spectromorphological* perspective, a method concerned with the depiction and impression of sound association (Blackburn, 2011:5). These would be necessary if no score was present, but is not to say their methods can't still be adopted, such as

the previously mentioned use of sonograms. In fact, Blackburn proposes the adoption of spectromorphology to create "informed compositional strategies" (Ibid). However, arguably the foremost expert on spectromorphology Denis Smalley, asserts that "spectromorphology is not a compositional theory or method, but a descriptive tool based on aural perception" (Smalley, 1997:107). The most beneficial use of spectromorphology in this project has been the adoption of the technique and applied somewhere between what the views of Blackburn and Smalley present, using the tool to make informed compositional decisions.

The disadvantages stem from the subjective perspective, where referential background may influence the outcome of the analysis, and where multiple objective views on the process may reveal hitherto unnoticed aesthetics and techniques present in the work.

As intimated, specific sections, particularly where procedural elements occur in tandem with the acoustic score, are given deserved attention. In sections where important data-driven events occur, these are also considered. In order to provide clarity in the text, each detection point studied will be written in square parentheses. For example [detection 1D], [detection 2P] and so on, where 'D' and 'P' denote data-driven or procedural events. The Sub-Patches that perform both forms of audio processing are presented in figures in each gesture sheet. For the parent audio patches, see Appendix II.

With regards to notation, the process of notating *In-Gest* approaches the syncretic, where original sketch material becomes electroacoustic material, generating acoustic relationships and gestures, which proceed to develop into both the acoustic and electroacoustic scores. Important to consider, as has been previously intimated, is that the score for the electroacoustic part is also the patches. Thus the electroacoustic score is a layered system of both modified traditional notation techniques and computer programming. Contextually, gestural mimesis between the two intrinsic scores has been detailed from a conceptual and technological perspective. In the sections that follow, examples from the sheets that demonstrate strong mimetic and gestural linking are shown with extracts from the scores with descriptions of the compositional techniques applied.

At the end of the general analysis, some examples of detection points relating to the GEM image data are shown to provide information as to how the variables for each detection point were determined.

6.2.3.1.1 - *Walk*

The intention in *Walk* is the primary exploration of the relationship between two concurrent musical gestures in each score that evoke a sense of movement as if slowly moving along at a steady pace. The violin opens with a lilting melody first with double-stops and then in single pitches which underscore the entire gesture sheet.

[detection 2P]:

The presence of acousmatic material in *Walk* was an important source of musical material and presented a particularly interesting notational challenge. The decision to include the representation of this element is motivated by the procedural events, where a 'random' walk is generated from timed events. The elements consist of layered acousmatic material of the gravel sounds underfoot. Extracted partials that relate to the original sketched material are built into two note chords, with their relative frequencies corresponding to those in the equally tempered chords. The thematic linking between the original motive and voices in the chords can be seen in the relationships between the bottom and top notes of each successive chord. Figure 24., below, gives an overview of the elements that are thematically linked.

Figure 24. - Thematic Linking in *Walk*

The figure is a diagram illustrating thematic linking in the piece *Walk*. It is divided into three main sections: A, B, and C.

- A. Original Phrase:** Shows a short musical phrase on a staff with notes G4, A4, B4, C5, B4, A4, G4.
- B. Acousmatic Material:** Shows two spectrograms of gravel sounds. The top one is labeled 'Heel' and the bottom one 'Toe'. Lines connect the 'Heel' spectrogram to the notes G4 and A4 of the original phrase, and the 'Toe' spectrogram to the notes B4 and C5. Below the spectrograms is a musical staff with notes G4, A4, B4, C5, B4, A4, G4, with 'nat.' written above it.
- C. Musical Notation:** Shows three musical staves. The top staff is for the right hand, starting with 'sul tasto' and 'p', followed by 'pp', and then 'mf espress.'. The middle staff is for the left hand, starting with 'nat.' and 'mp molto legato'. The bottom staff is for the right hand, starting with 'mp legato'.

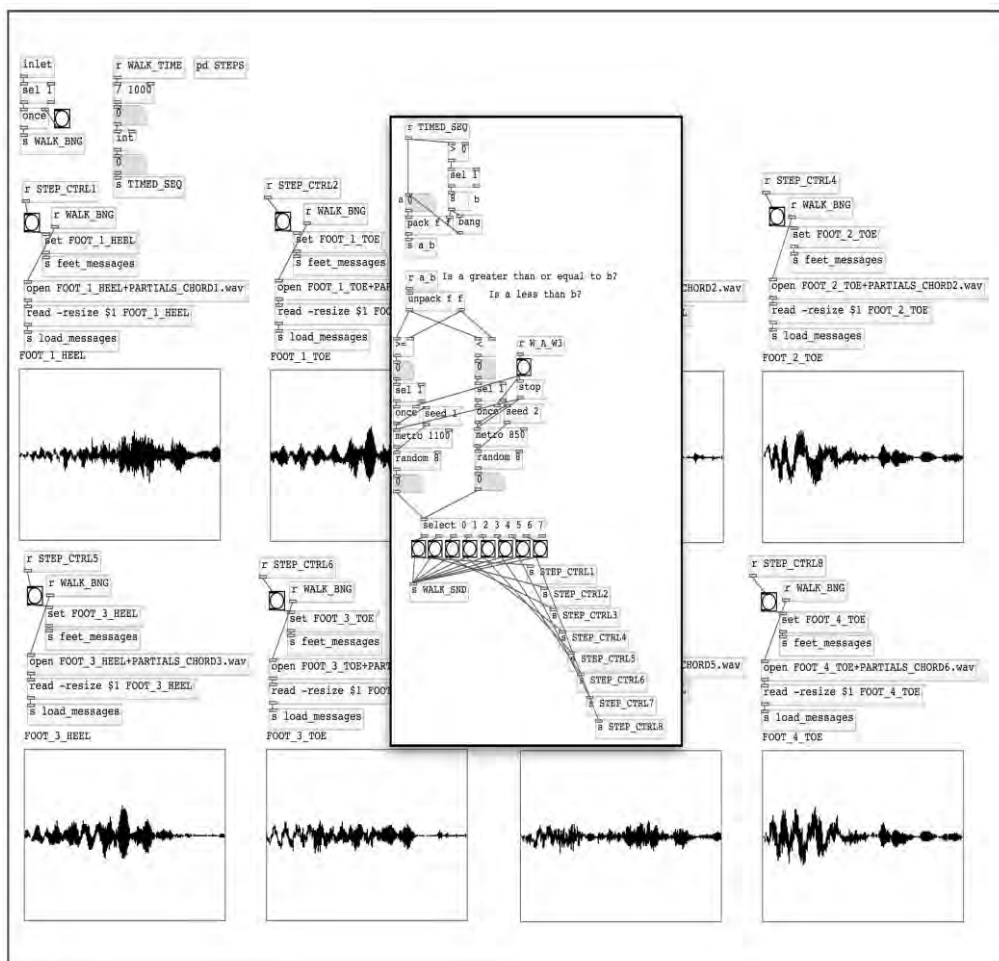
In C. where the electroacoustic part assumes the leading role, the chords from the acousmatic material are built into vertical clusters of between three and six pitches. The intention here is to gradually create a sense of density prior to the rarefication of the harmonic material that undergoes pitch-shifting until a decrescendo to silence. The gesture ends with a mixture of the two note chords in pizzicato followed by single beat interjections of the harmonic material. The procedural material at [detection 2P] is notated by showing a sonogram cross-section (B.) of the original material (A.) relating to either the heel or toe of a particular step.

This cross-section of the sonogram shows the extracted partials that relate to the chords in the acoustic part with a conventional representation of the chords in the final part of the image. The sounds underwent analysis and filtering to extract specific harmonics and partials that relate to the first eight chords of the composed acoustic sequence in the violin part. The sounds are layered and placed in the patch to be played back randomly depending on the timing between the first two detections. Figure 25., overleaf, shows the PD patches as developed which enable the timing of the events as indicated in the score to control the randomisation of the sounds, as if each performance gives a different 'walk'. [detection 3D] triggers the sequence to end.

[detection 4D]:

The score at this point results in the electronic part having a solo role after the last note in the violin. The chords heard are layered harmonics extracted from the samples as per Table 4., with the further addition of some harmonic details from the [detection 2D] material. This creates an assimilated sound between the two scores, where there are common timbres but occurring in differing contexts. Furthermore, the relationship with the acoustic score is kept close where the structure of the chords reflect the last sequence of pitches (C, G#, A, B).

Figure 25. - *Walk* [detection 2P] Sub-Patches



[detection 5D]:

The electroacoustic score here makes use of the recorded harmonic material from the acoustic sketches as per Table 4. The notated harmonics are formed from the adaptation of the opening melody. The electronic score has transformed material where each sequence of the harmonics are pitch-adjusted by either a $\frac{3}{4}$ or $\frac{1}{4}$ tone as if the harmony of the entire section becomes more rarified until it disappears in silence with just the harmonic C# left on the violin. The section is followed by a call-and-response motive between the original chords in pizzicato and bowed harmonics, before concluding with the final five chords of the sequence.

6.2.3.1.2 - *Up*

The first impression of viewing the score in *Up*, is one of notated sonograms. This is certainly the intention here, where material generated is layered and then assembled in large vertical clusters with time-stretching applied and then analysed in the software package Sonic Visualiser⁵⁹ to produce the sonograms. The sonogram of the audio material created and used at [detection 1D] in the first section is shown in Figure 26., overleaf. Noticeable is the overall direction of the frequencies, which are upward-trending. This is supported notationally in the electroacoustic score by the design of a symbol that indicates a raised

⁵⁹ See <http://www.sonicvisualiser.org> for details (accessed 2014, June 15).


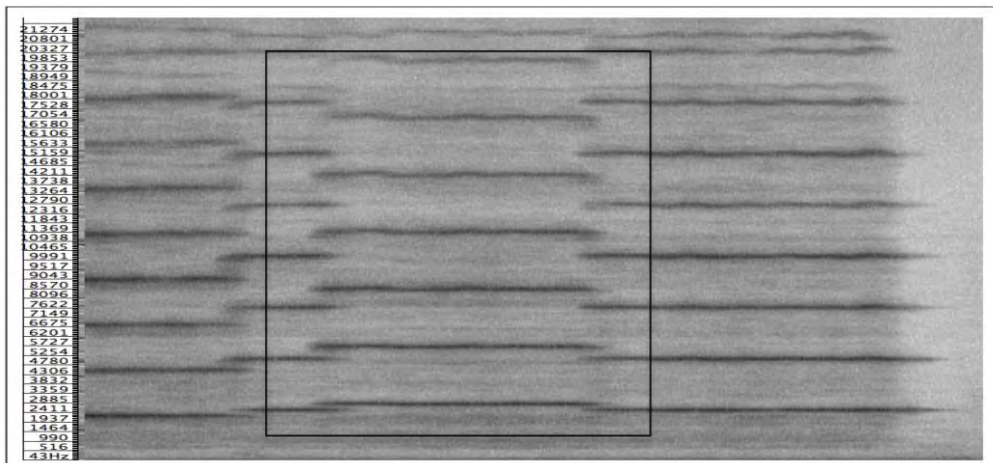
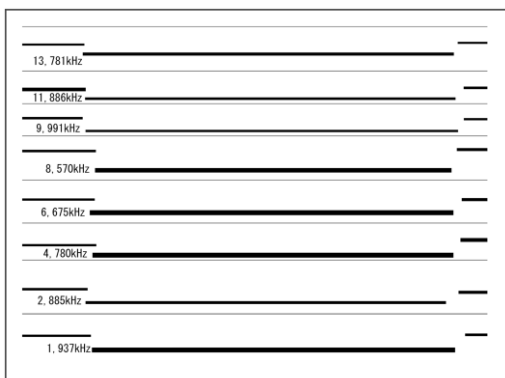
pitch (#), with a general trend towards sharpening when applied temporally: 

Figure 26. - *Up* First Section Sonogram



In the figure below of the notational elements, one can see a clear link between the image of the sonogram and that of the notation in a technique borrowed from Karlheinz Stockhausen's (1928–2007) *Telemusik* (1966) (Dimpker, 2013:214). The square outline from Figure 26., is the section extracted.

Figure 27. - Section from [detection 1D] Sonogram Notation



The relative thickness of the horizontal lines indicates an approximate relational amplitude level of each of the frequencies.

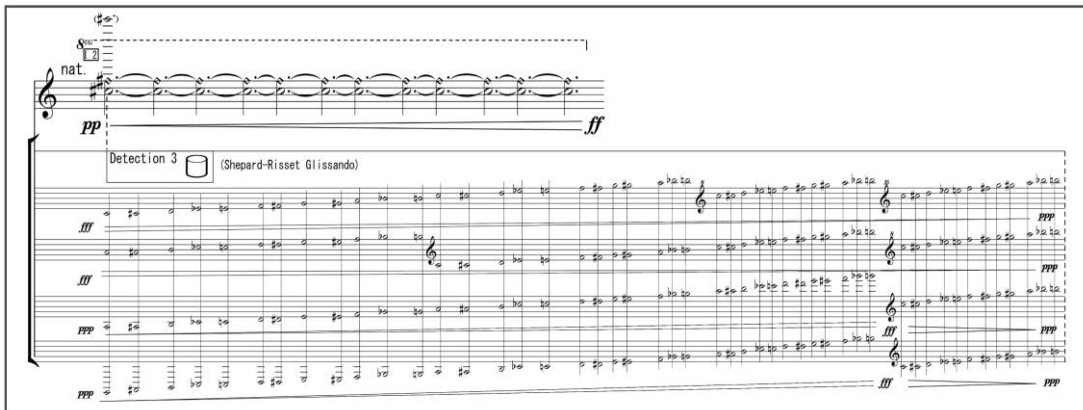
[detection 2P]:

At this point in the score, the material is procedurally transformed with finger-tracking data collected from [detection 1D] until [detection 2P]. Depending on the variables in the data (x;y), the transformations applied are either an increase or slight decrease in pitch, with the intention that mostly an increase in pitch is applied by limiting the input range of the (x;y) variables. The material is related to that from [detection 1D], in that it is similarly constructed, but with a marginal increase in tuning by a few cents and the addition of a further octave transposition.

[detection 3D]:

The audio that is layered with the acoustic part here is transformed to include an auditory illusion of an infinitely upward sliding scale. The construction is formulated from initial experiments by Roger Shepard, who created the illusion of an endless increase in pitch by adjusting "sinusoidal components" a semitone each time and subsequently extended by composer Jean-Claude Risset with a glissando technique (Pierce, 1999:67,68). Notation for this section required the spreading of the octave transpositions of the sliding scale over the length of the glissando with the critical dynamics as applied. Figure 28., overleaf, shows how this was achieved.

Figure 28. – Notation of Shepard-Risset Glissando at [detection 3D]



The acoustic score is gradually increasing in pitch position with the application of $\frac{1}{4}$ and $\frac{3}{4}$ sharps between larger elements that are trending towards a higher register. These rudimentary phrases (elements) are built-up from several variations, bound by the missing quavers from their dotted crotchet groupings. Figure 29., on the next page, shows these phrases as constructed. The relationship between the two scores is now well established, since they are both mimetically and gesturally linked from both a musical and instrumental gesture perspective. The tracking events of the performer's fingers during the interpretation of the section between [detection 1D] and [detection 2P], have a direct influence on the outcome of the musical material from [detection 2P] until [detection 3D]. A variable slider is controlled by the (x) parameters from the tracking, which in turn adjusts the pitch in half-tone transpositions. This is achieved with the G09.pitchshift.pd abstraction available in the PD browser. In essence this is an emulation of a classic rotating-tape-head

machine. Figure 30., overleaf, shows the sub-patches responsible for this procedural manipulation.

The penultimate note in the violin part after the last phrase variation's quaver rest is a C $\frac{3}{4}$ sharp. This intentionally 'leads' upwards to the final jump to the last section at [detection 3D] where the gesture ends with an artificial harmonic on the highest note (C#) a violin can produce.

Figure 29. - Phrase Variance in *Up*

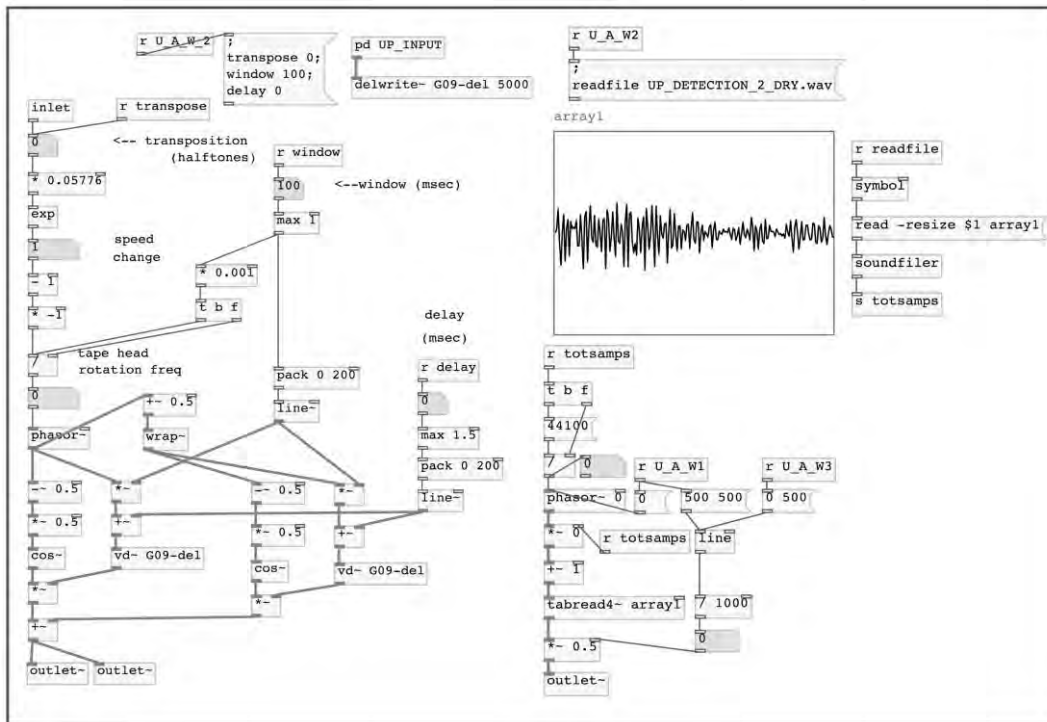
The image displays a musical score for violin, consisting of four staves. The first staff begins with a dynamic marking of *mp* and includes a 'vib.' (vibrato) instruction. It features a series of eighth notes with slurs and accents, leading to a final note marked with a '7' (fingerings). The second staff continues with similar eighth-note patterns, also marked *mp*. The third staff shows a transition to a more melodic line with slurs and accents, marked *mp*. The fourth staff starts with a dynamic marking of *f*, followed by a section marked *f* and *mf*, and concludes with a final note marked with a '7'. Above the staves, there are annotations including '8va' (octave up) and a circled '8' in the fourth staff, indicating specific performance techniques or pitch bends.

6. 2. 3. 1. 3 - *Down*

Anomalous Low Frequency (ALF's) or 'Sub-harmonics' is an extended technique for the violin thought to have first appeared in George Crumb's (1929) *Black Angels* (1970) (Strange, 2001:24). The technique has subsequently

been developed, perfected and used extensively by performer and composer Mari Kimura (Kimura & Risset, 2006:408).

Figure 30. – *Up* [detection 2P] Sub-Patches



It is applied here where initially a recording was made of a short sub-harmonic on the violin and transformed as per Table 4. The ALF's produced were isolated and then enhanced by applying amplitude gain to the low frequencies. Furthermore, in some instances over-driven distortion filters applied to the sounds provide impetus to the already 'grinding' sound. The overall objective here is that a sense of descent is created through a technique of mimesis between both scores.

[detection 3P]:

There are several elements contributing to the overall sound design at this point. Firstly, the initial sample from Table 4., has been time-stretched slightly and then cut into many sections of differing lengths as if cropping a musical phrase into small sections and then changing the durations of the pitches in each section. These were then sequenced into randomly selected positions based on the constantly changing harmonic rhythm of the original ALF. The procedural patch at this point determines the timing between [detection 2D] and [detection 3P] and then applies the result to a method of randomisation where the timing between events constantly changes as well as the order. The musical gesture implied by this is one of 'falling' down an uneven slope. Figure 32., overleaf, reveals the inner workings of the patch creating the randomised sounds.

There are essentially two 'pitches' of ALS frequencies. The first is the creation of a low D sub-harmonic by applying extreme pressure at the 3rd partial node on the G string. The other is applying similar pressure over the node where the player would usually perform an open low G. They are indicated by square note heads with their relative sub-harmonic 'pitches' indicated in an ossia score below. Perhaps the most important element besides the sub-harmonic pitches is the rhythmic mimicry taking place. In the last section of the opening block, the electroacoustic score mimics the acoustic score rhythmic gesture of the ALS's and then reverses it, where the gap between repeated notes gets slightly larger each time in a

form of augmentation. The rhythmic motive is also layered with each layer of the motive proceeding in opposite directions. The decelerating *saltato* element at the end of the acoustic score is a final 'trip' before the last descent into the depths. Figure 31., below, places the two elements together to highlight the rhythmic linking.

Figure 31. - Rhythmic Linking Between [detection 2D & 4D]

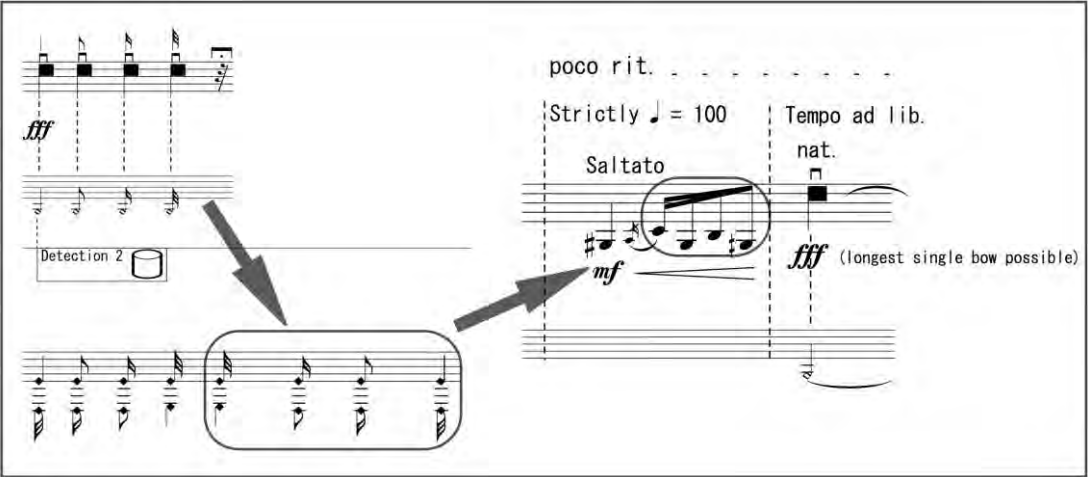
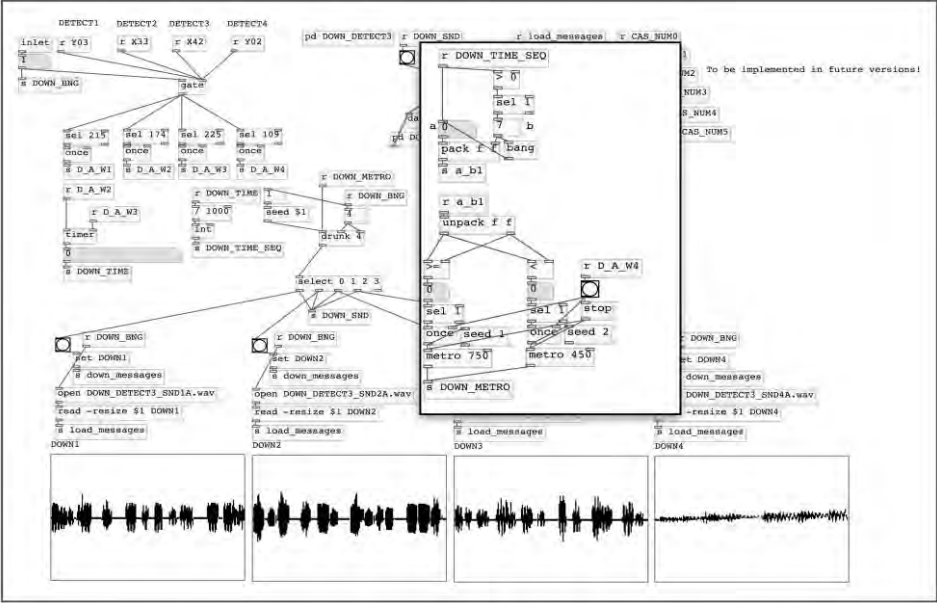


Figure 32. - Down [detection 3P] Sub-Patches



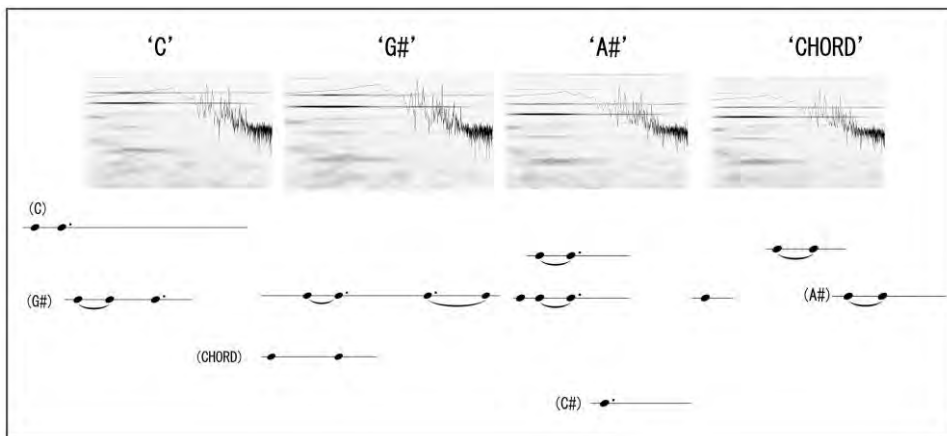
6.2.3.1.4 - *Waltz*

The system created to perform the score of *Waltz* is perhaps one of the more complex ones in the overall work. This is because a real-time envelope creation tool was designed to allow the finger tracking information with six simultaneous co-ordinates being streamed, to control amplitude envelope parameters and then apply them to the pre-recorded and transformed sounds. The desired inclusive effect is one of quickly fading and morphing musical gestures as if the source of the sound is whirling around a dance-floor in the performance of a choreographed waltz. The choreography here is the acoustic score and the careful placement of the detection points. Where the other patches mostly rely on timing between points, this patch relies on the finger-tip tracking data entirely, especially the constantly changing (x;y) parameters. Thus, each of [detection 7,8,11 & 13P] points performs the same function, except the envelope parameters will be very different each time.

The material captured and generated for this sheet consists of single pitches isolated from the original sketches. Layering of the single pitches provided the sounds as indicated in the score. The 'chord' is a combination of G#, A#, C played as a cluster on the D and A strings in the generation phase as per Table 4. The notation shows a single pitch, with its root note written in text, but it symbolises a harmonic structure that was determined using a combination of spectra and partial detail analysis.

In a way, this is a permutation of the technique adopted by Cogan (1984:9,10), where each of the “*sounding elements (that), together, make any sonic impression.*” A comparison of the sound event’s resulting images enabled the determination of which ones should be used. Figure 33., below, includes the combined analysis techniques of spectra and partial extraction together with the chords as applied in the score. Noticeable are the similarities in the contours of the sounds.

Figure 33. – Electroacoustic Chord Structure in *Waltz*



The attack and decay of each sound as shown above (besides 'CHORD'), was altered and then equalisation and filtering applied to remove the intrinsic vibrating string sound of an acoustic violin. What is left behind is something that sounds like the striking of strings by a hammer - i.e. piano, instead of a bowed string, but with most of the attack missing. This could be viewed as a loose analog to the tradition of salon-style violin and piano ensembles of the late 19th century, where the waltz as a musical style strongly featured. Figure 36., page 159, gives an

inside look at the sub-patches responsible for the envelope generation and application.

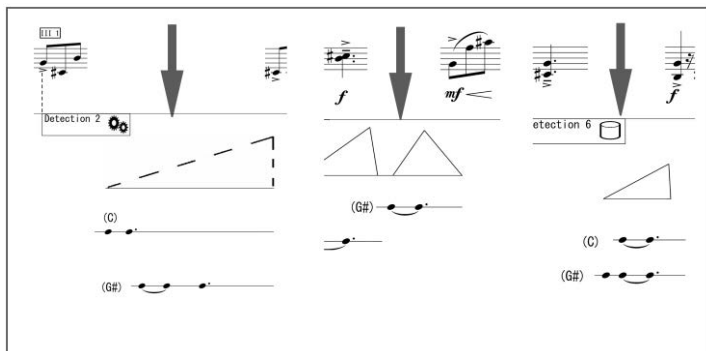
From a notational perspective, the two scores are related through timing and placement of events. There are potent rhythmic and dynamic superimposed elements taking place. In the acoustic score, the three-quaver motive undergoes extreme variance, with each occurrence perceived as a dotted crotchet in totality. The figure below gives a breakdown of the variations.

Figure 34. – Rhythmic Variance in *Waltz*



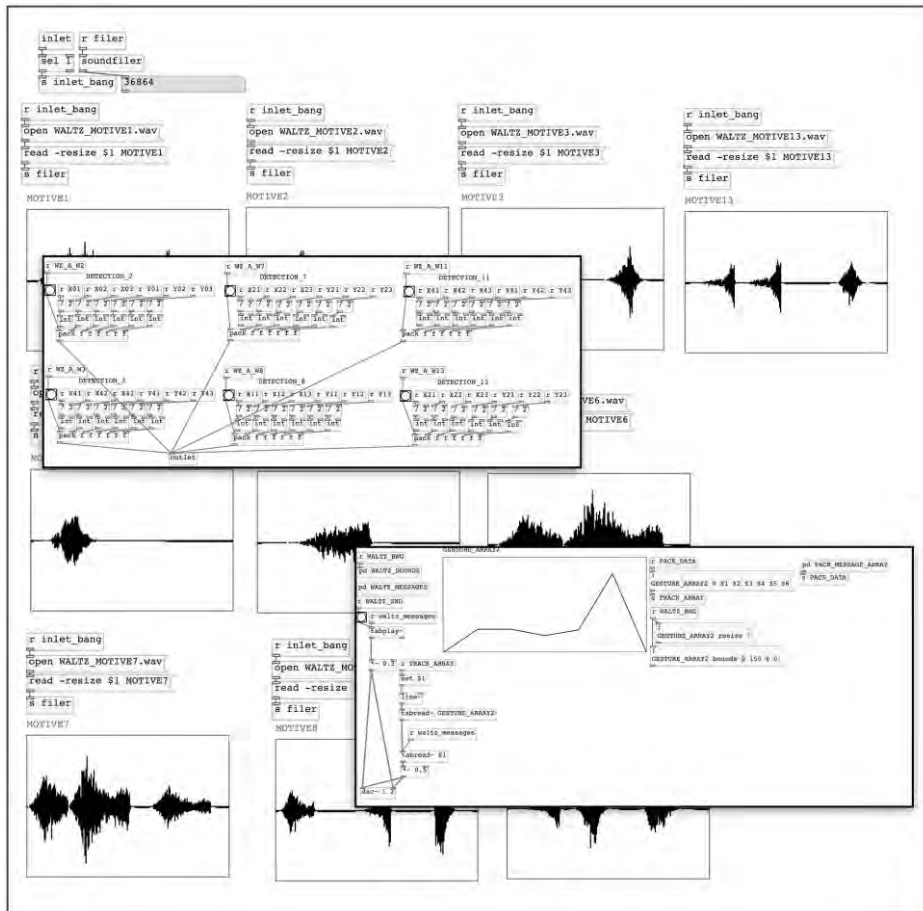
In a comparable vein where the envelopes are applied and generated through procedural processes, the electroacoustic score is provided with dynamic variance and the placement of the detection points provides the acoustic score with motivic interjections. This creates an antiphonal effect. An example of this is shown in Figure 35., on the next page.

Figure 35. – Motivic Interjections in *Waltz*



The dotted lines of the envelopes indicate positions where procedural envelope information is applied to the motives that are provided data sets from the finger tracking information. Where the envelopes are intended to be fixed, they are indicated in the score with solid lines and applied to the data-driven events as set elements. Consequently there are two sets of audio material for the electroacoustic score: the one set has pre-applied envelopes, the other none, waiting to have them applied in live performance.

Figure 36. - *Waltz* [detection 7,8,11 & 13P] Sub-Patches



6.2.3.1.5 - *Float*

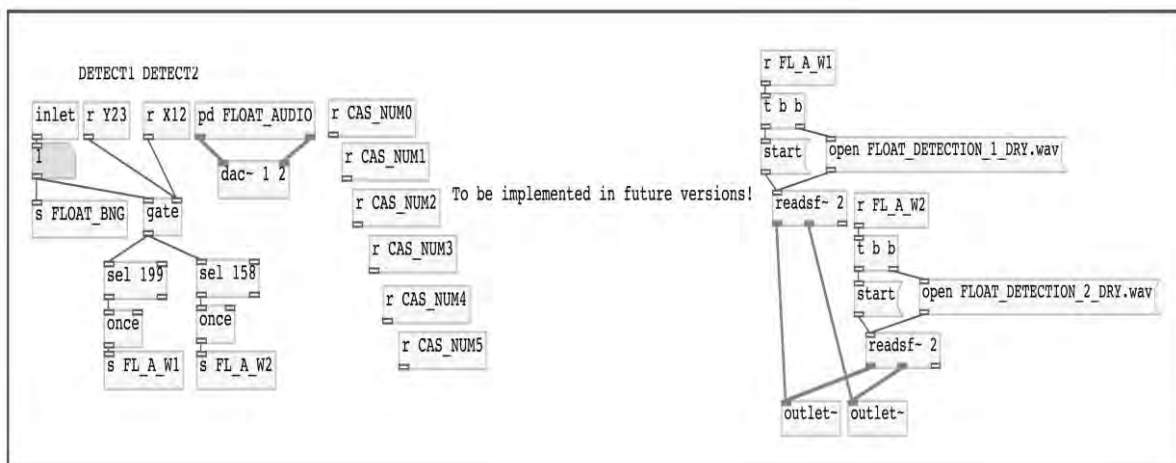
In a digression from the normative where each sheet has procedural events, *Float* is the exception to the rule. Only data-driven events occur in this part of the work. The compositional reason for this stems from the necessity to provide some baseline structure to a musical gesture that is particularly 'loose'. Floating evokes images of transcendence, lightness and rarified atmosphere. The extremely open nature of the acoustic score where no strict tempo is given (tempo ad lib.), provides the performer with tremendous carte blanche. Whilst the electroacoustic elements have

fixed lengths (as indicated in the score by time ratios), these are not intended to influence the performer nor vice versa, but rather to act as musical comments, interjecting themselves between the various violin motives.

The material used in this sheet mostly consists of layering to provide chord structures with multiple octave transpositions applied. The detached and somewhat distant sound is achieved by using extreme time-stretching techniques on very short motives extracted from the chord structures but without changing the pitch. The only pitch changes occur from within the motives themselves, such as the bi-directional glissando at [detection 2D].

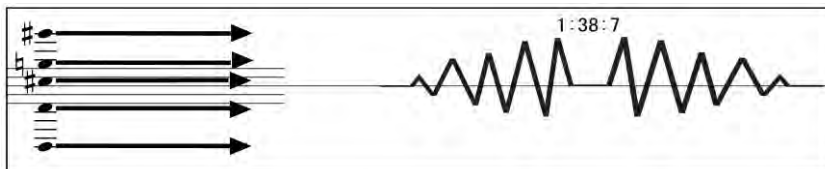
Figure 37., below, shows both of the patches created to realise the complete musical gesture in performance.

Figure 37. - *Float* [detection 1 & 2D] Main and Sub-Patch



Notationally, the acoustic part appears to be far more complex than the electroacoustic. However, the notational innovations of the electroacoustic score draw inspiration from Ferneyhough's (1943) *Time and Motion Study II* (1978). In the Ferneyhough, lines that become wavy and then straight again indicate variance or time-shift in a delay. Here a similar notational device is employed, only the lines indicate the start and end of a wave-form sequence, delineated with its overall length in a time-stamp ratio. The figure below, shows an example with a chord stretching out over a duration of 1 minute, 38 seconds and 7 split seconds. The method is also applied in other gesture sheets where necessary.

Figure 38. - Wave-Form Time



From a motivic perspective, there are four important gestural elements taking place in this sheet. They are:

- *flautando* effects;
- high trilled glissandi;
- extended technique of bow screw bouncing on the strings;
- high tremolo double-stops.

They are related from a pitch point-of-view and are rhythmically all ad libitum. Apart from the bouncing screw sound, their timbre exhibits similar characteristics that evoke lightness and transparency in line with the overall aesthetic intention of the sheet. The effect created by the screw bouncing is one of very soft short metallic bursts. They are meant to be loose acoustic resemblances to the sound effects created by Iannis Xenakis (1922–2001) in *Concret Ph* (1958) that appear to *float* in the space they were composed for, where “music, sound and space are the same thing for the listener” (Meric, 2011:1).

Besides being harmonically linked to the electroacoustic score with the layering of octave transpositions as discussed, the stretching of the original material is an important structural transformation that echoes the free-form rhythm of the acoustic part. This is particularly noticeable in the section where the bi-directional glissando is time-stretched to over two minutes in length in a high transposition. The result is that the location of pitched material in vertical relationship between the acoustic and electroacoustic score is entirely dependent on the tempo at which the performer decides to take the music. The intention is to create ephemeral shifting of harmonies that shimmer in constantly changing modes of consonance and dissonance, which in turn become an outside-time structure⁶⁰

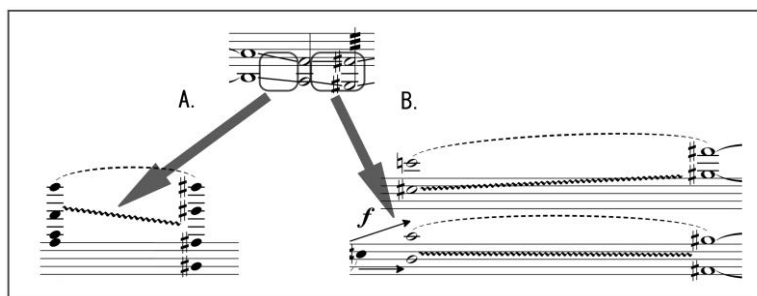
⁶⁰ *Outside-time structure* refers to musical events that are organised independently of their occurrence in the music as it unfolds in linear time.

compositional feature. The section marked "zephyr-like" is in homage to Crumb's *Black Angels* (1970).

6.2.3.1.6 - *Slow*

The single most important overall element present in *Slow*, is the extended glissandi in both scores. Constantly surrounded by long sustained motives, each of the glissandi in the acoustic score must be performed at the slowest possible rate. Where they are simulated in the electroacoustic score, the glissandi start at a specific tempo and then gradually decrease in speed until the next sustained pitch cluster is reached. The implied gesture here is a sense of slackening towards dissolution. The image below shows the thematic and gestural linking of these elements.

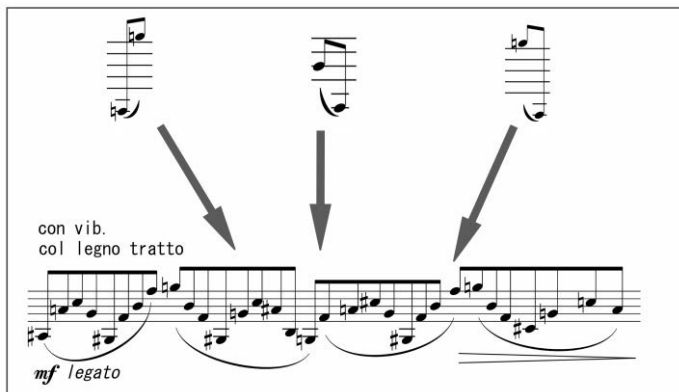
Figure 39. – Thematic and Gestural Linking in *Slow*



In A., the downward mono-directional glissando is mimicked, whilst in B., the bi-directional glissando undergoes the same process in the electroacoustic part.

Moving further along in the score, at [detection 4D], the *col legno tratto* section in the acoustic score is a hyper-augmented setting of the short appoggiaturas present as short breaks in the long drawn-out gestures of the opening part. The compositional apparatus here is a further example of augmentation that creates a sense of slowing down by increasing the length of the smaller elements. This is accompanied by a rapid and time-stretched tremolo on A# in the electroacoustic part, superimposed on a set of minor 7ths. Figure 40., below, presents the augmentation of the motives as applied.

Figure 40. – Augmentation of Appoggiatura Elements



All but one of the detection points are data-driven, where the elements from Table 4. have been applied in various transformations. All the chord structures are constructed from the ubiquitous minor 7ths present in the score, which took form from the initial sketches as shown in 'Gesture I' of the table.

[detection 3P]:

With the chord motive continuing in part of the musical material present at this point, there is the addition of a procedural element, which is the live time-stretching by granular synthesis of an extracted partial element from the generated materials. Granular synthesis is the technique of “creating complex sounds from tiny units of sound.” (Helmuth, 2006:189). The partial motive is manipulated based on the timing the violinist takes between [detection 2D] and [detection 3P]. Thus, should the timing be X, the rate of playback per grain is Y and so on. Figure 41., on the next page, gives a representation of the granular synthesis patch⁶¹ as applied. The notation of the procedural element here required the creation of a graphic to represent the partial extracted together with the application in context of the waveform graphic, but without a time-ratio since the length of the element is indeterminate. The graphic in Figure 42., overleaf, shows the notational device in context.

6.2.3.1.7 - *Fast*

In comparable complexity with the patches developed for *Waltz*, many parallel processes are taking place in the programming behind the electroacoustic score in this sheet. To provide clarity, they are annotated below:

⁶¹ This sub-patch (with additions added where necessary), was garnered from <http://www.pdpatchrepo.info/hurlleur/timestretch2.mmb.pd> (accessed 2014, August 24).

- [detection 2,3,5,6,8,9 & 10P] are providing a timer with triggers to start and stop timing, out-putting logical time in milliseconds at each subsequent detection;

Figure 41. - *S/ow* [detection 3P] Sub-Patch

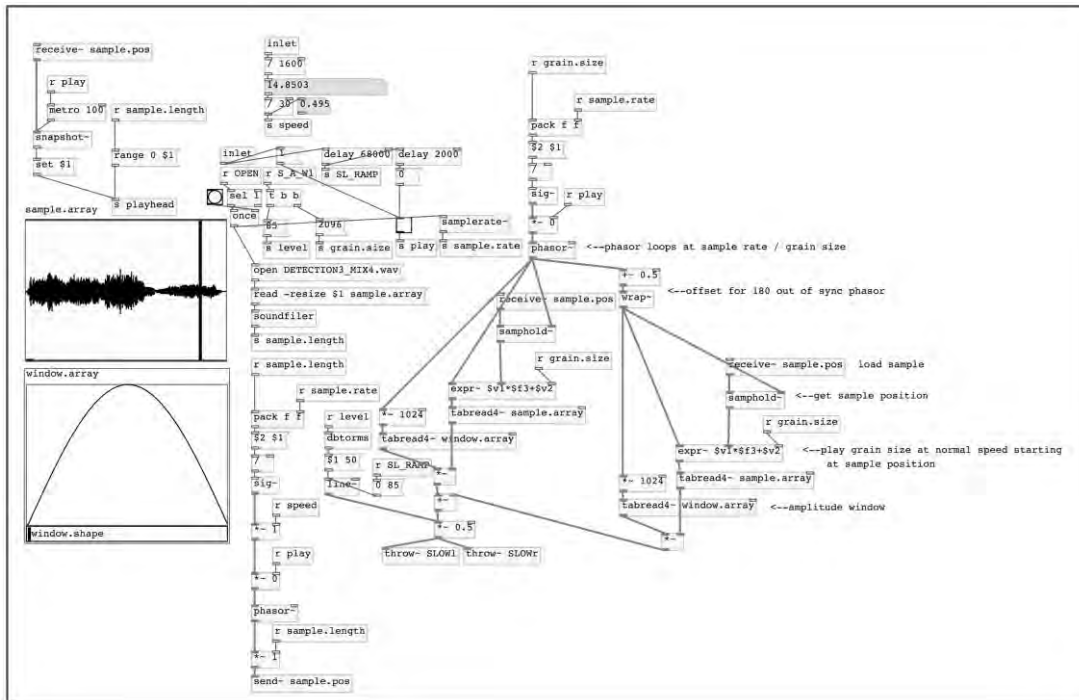
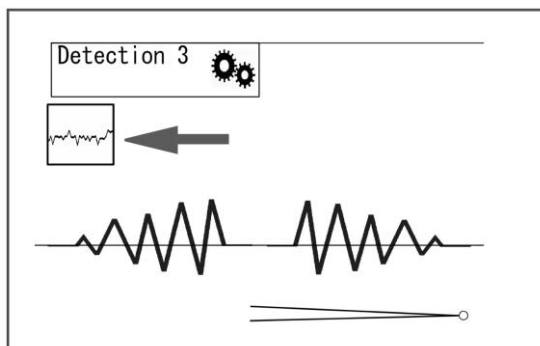


Figure 42. - *S/ow* [detection 3P] Notational Elements

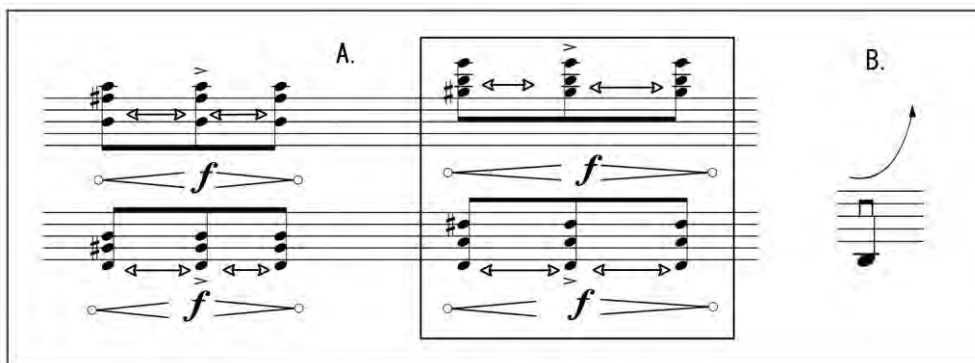


- the timing data is used to feed information into the patch that controls the playback rate of the audio data;
- logical operators control the route of the timing data prior to being applied to the audio information.

Each of the audio samples delivering material for the patch contain several transformations as applied to the original material generated. Layering created chords that could be used as individual elements. With these elements assigned in pairs to form units, an important transformation by means of the reversal of audio content was applied. This is then further speeded up in the procedural patch, assisting in creating an artificial Doppler effect. Whilst Doppler effects are created in simulation by "variable delay" (Settel, 2001:92), here the time-stretching application of the original sample with additional reversal applied to the two audio events joined back-to-back mimics the delay and provides a "sense of motion of a sound source" (Pierce, 1999:97). The acoustic score contains a technique where the player is requested to accelerate the bow without changing the dynamic. The artificial Doppler effect in the electroacoustic part is mimicking this element. Thus they are intertextually linked, providing musical gesture affordances between each other. Figure 43., part A., overleaf, shows the notational techniques developed to indicate this effect. The arrows signify the implied variation in speed of the material. For example, in the first element on the left-hand side of the figure, the top chord (B, F#, A) has

acceleration and deceleration taking place simultaneously, which is perceived as one chord not three separate chords. The element on the right is boxed and therefore undergoes procedural manipulation. Each of these motives has crescendi and decrescendi starting and ending in silence respectively. These motives are gesturally linked with the bowing technique of acceleration without changing dynamic, as indicated by the upward curving arrow in part B. of the graphic. Figure 44., on the following page, provides a captured image of the sub-patches responsible for the processing.

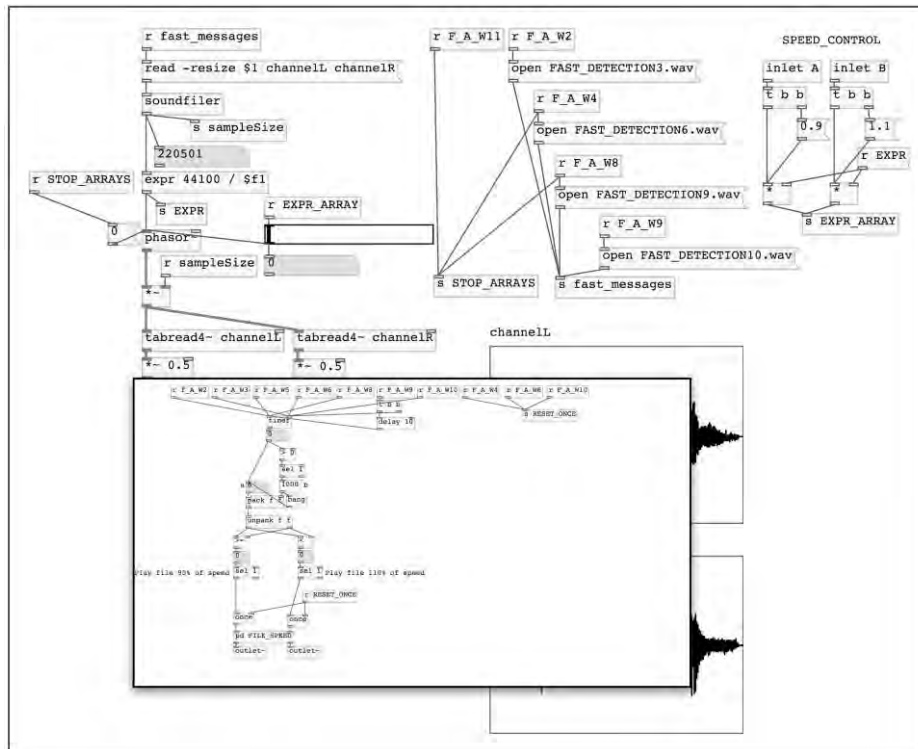
Figure 43. - Notation of 'Doppler' Effect



6.2.3.2 - Determination of Tracking Data

This section is included at this point in the thesis, due to the requirement that the scores be complete, particularly the acoustic score

Figure 44. - *Fast* [detection 2, 3, 5, 6, 8, 9 & 10P] Sub-Patches



before the tracking data could be used in the patches. Throughout the development of the patches supporting the function of the heuristic as applied, images were captured of the finger-tracking points and aligned to the score where detection is required.

Each detection point in the score is accompanied by two types of symbols: 1., the detections themselves such as [detection 2D] etc., and 2., a boxed set of Roman numerals and a number:

III	3
IV	1

 This symbol indicates the exact fingering that must be adopted at that point in the score. The Roman numeral indicates which string must be used and similarly, the numeral which finger. If two sets of numbers occur as in the example shown, then they correspond to a chord where the fingering applied is from the top

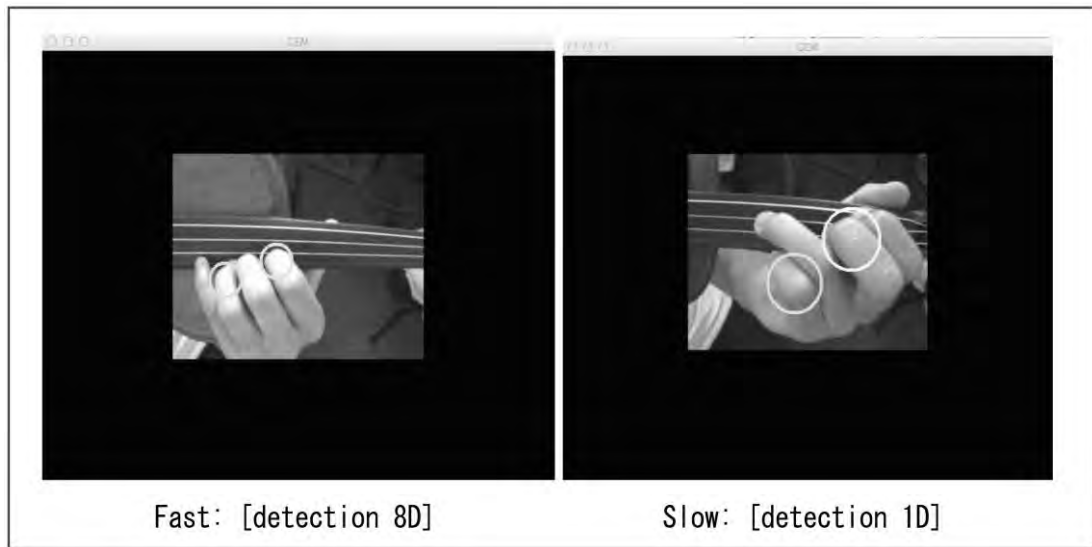
note of the chord moving down. This helps to provide static object tracking data to the system.

The figures that follow show some examples of the data captured and how this is applied in the patches concerned.

In the first image (Figure 45., overleaf), we see two separate examples of what the GEM output reveals at the specific score locations as indicated. This means that even fingers that are not directly involved in producing sound also provide tracking information if the system detects them at that position. To limit the wide variables of tracking data, programmed ranges of tracking data are specified to be processed per detection point before sending a specific trigger. A set of sub-patches to control this through a method of data selection provided this necessary regulation.

Figure 46., on page 172, shows the sub-patch that contains the individual patches relating to each gesture sheet and an example of one of the embedded sub-patches that contains the objects for setting variable ranges. A brief explanation of its processes follows.

Figure 45. - GEM Detection Point Image Capture

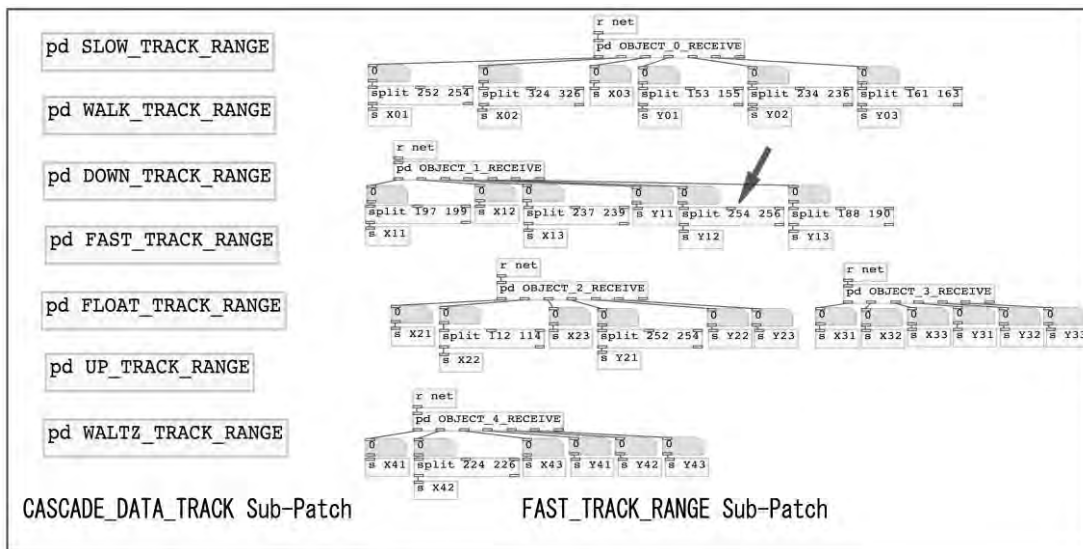


The [split] object contains a range of pre-set arguments such that if it is [split 200 202] as in the extract from [pd WALK_TRACK_RANGE], it will only allow output of the range 200-202 (exclusive) as integer streams. Thus, the integers in-between are output as running integer streams. Integers outside of the range are only output as either maximum or minimum as per the initialised argument. Critically, this means that any other tracking data being received as a constant stream from the VIDEO feed will not interfere with the other detection points in the patch.

The range argument variables were determined by examining the initial capture data as per the example in Figure 45., for each detection point and then added. It may happen that these variables could change from violinist to violinist, performance situation to performance situation,

and would require a run-through with the controller to perhaps alter the [sel XYZ] variables for each detection point as needed.

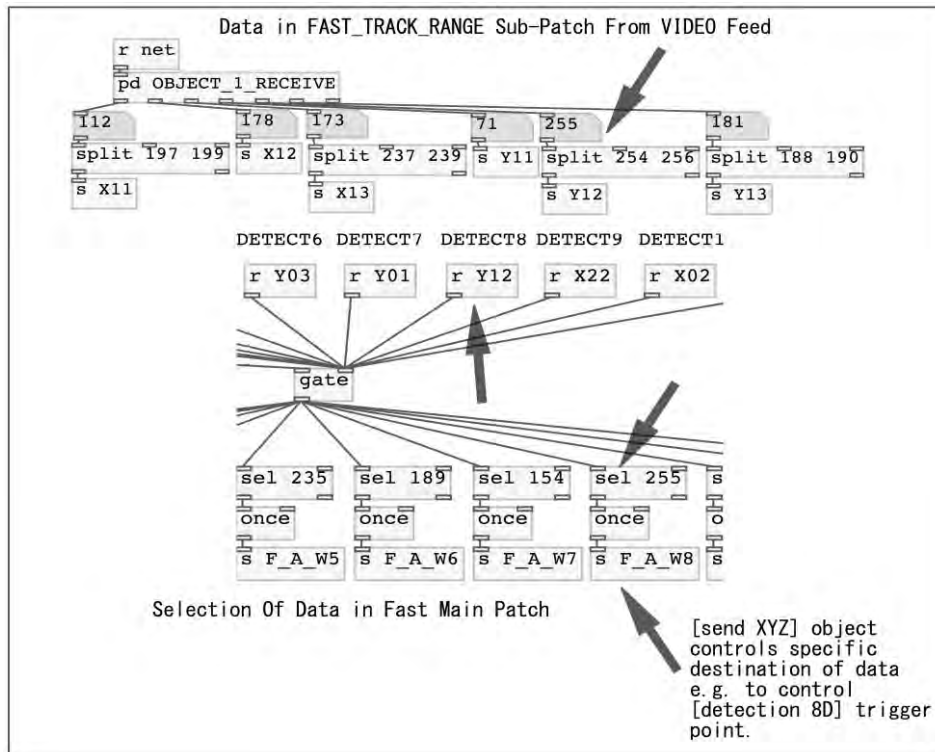
Figure 46. – Data Range Control Sub-Patches



A future enhancement could be the addition of an easy-access panel to enter the new selection data per detection point, and an additional improvement would be if the object identifier number coming from the [pix_opencv_haarcascade] PD object could be used without the system dropping the connection as discussed in section 1.4 of Chapter 5. Overleaf, in Figure 47., the data is correlated finally in each patch. Following the arrow from Figure 46., the information is gathered in the VIDEO patch and sent to the CASCADE_DATA_TRACK sub-patch for range-setting as discussed and then routed to each gesture patch. This distributes the data per object in (X1; X2; X3; Y1; Y2; Y3) format and finally this is selected with the [sel XYZ] object in each gesture patch, which allows

only the variable XYZ through. Ultimately this is patched to the intended destination with a specific [send XYZ] object, which has a partner [receive XYZ] object.

Figure 47. - Finger Data Correlation for *Fast* [detection 8D]



6.2.4 - Synopsis

The section provided here gives a complete overview of the compositional process as applied and approached in the creation of *In-Gest*. These are considered within the framework of a technical appraisal, whereas a discussion of aesthetics follows in section 3 of this chapter.

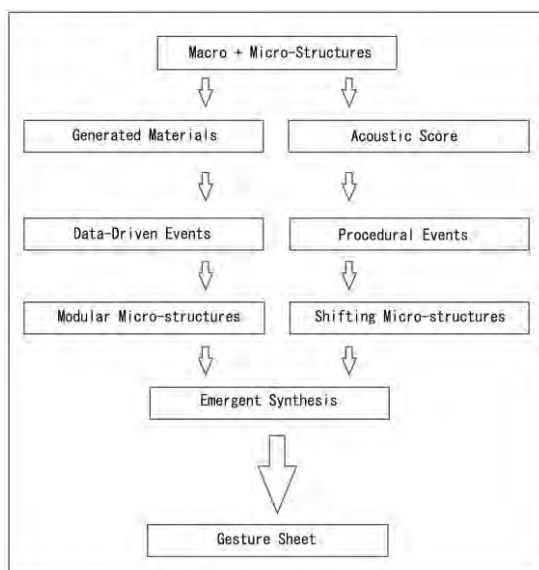
In much the same way that Boehmer (2002:66) describes Gottfried Michael Koenig's (1926) *Essay* (1957) as comprising of sections where the "quality of the form was defined by criteria that initially concerned the material", the complete macro-structure of *In-Gest* exhibits a similar approach. Materials gathered from the composition of small micro-elemental phrases generated electroacoustic material, which later informed the complete score. Furthermore, the aleatoric possibilities that exist in the macro-structure, provided by the possibility for each sheet to be performed in any order, create free narrative forms from structured elements. Thus the 'larger constructions' present themselves with each new performance as the performer chooses narrative. This is very similar to how the Cologne School of electronic music approached the conception of serial forms as "modular" consisting of small unchanging elements that form larger more aleatoric constructions (Boehmer, 2002:63).

Significantly, in turn, each of the gesture sheets which form the macro-structures of the complete work, are themselves constructed of 'modular' elements as explained. Thus the macro-structures are themselves containers of fixed events, since the underlying electroacoustic material has already been generated. However, there are the procedural events, which use fixed material to generate shifting micro-structures in each sheet. These processes provide the work with an *emergent synthesis* in each gesture sheet (apart from perhaps *Float*, although the acoustic score is very free), as if it forms its own new musical constituents from the generated

materials obtained in the first stage of the work's development. The description by Hamman (2002:107), where an "emergent model of musical material" that "allows the composer to find descriptions and criteria that arise...from the particularity of (her) activity", applies. See Figure 48., below, for an outline of the processes contingent with the *emergent synthesis* and its context in the compositional process in *In-Gest*.

With a similar methodology applied, the individual gesture sheets can be represented by images, with each image in turn represented by sonic forms. If these sonic forms embody the images through a sound-tracing process⁶² (where an individual listens and outlines the structure of the sound in a drawing/trace), then they are what Godøy (2006:149) describes as "*Gestural Sonorous Objects*".

Figure 48. – Emergent Synthesis in *In-Gest*



⁶² See section 2.1.1 of Chapter 3.

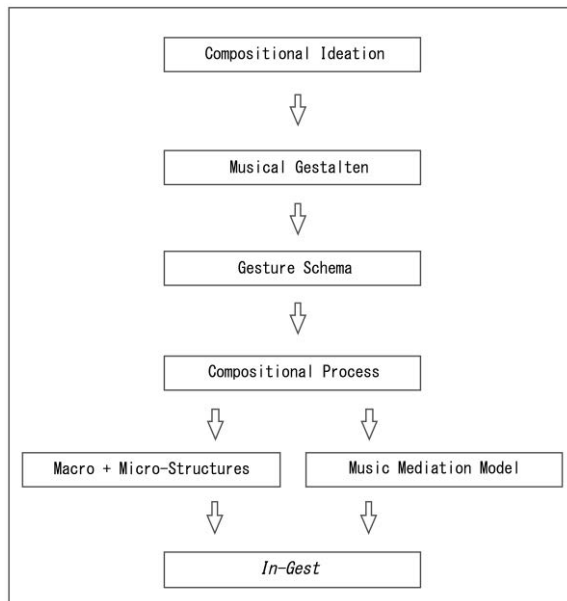
With the possibility for the application of this in composition, the tracing of these sonic images can be represented in diagrammatic form. Much like the “directions of movement” detailed in the analysis of Koenig’s *Essay* (Boehmer, 2002:66), the figure below, utilises a similar technique in showing the implied movement as imagined for each sheet of *In-Gest*.

Figure 49. – Directions of Movement in *In-Gest*



The work’s overall structure can be summarised by applying a conceptual apparatus in the form of a *structural scheme* that took form from the hermeneutic processes as already discussed. Figure 50., overleaf, provides a top-down view of the entire creation process as subjectively perceived.

Figure 50. – Structural Scheme of *In-Gest*



The imperative of technology in the process as outlined in the image becomes clear here. A technology that when applied in a heuristic sense, and developed in tandem with a musical poiesis, has the ability to act as an indispensable cog in the wheel that is compositional process. Where Mazzoli (2014:24) states that electroacoustic composers adopt the process of treating musical *Gestalten* into sonic images, the practice of composition that use, development and implement mediation forms to realise these images, provides those technological forms contingent with musical creation, an important place in the context of musical art forms as a whole.

6.3 - Aesthetics: A Consideration

The signification of the musical idea in *In-Gest* comes from the premise that structured sound is viewed as a sign – a conveyor of meaning. We have seen this through the proposed gesture signification models and how the hermeneutics of the two scores and their micro and macro-structures build upon the overall semantics of each sheet and ultimately the narrative from performer input. Demers (2010:23) argues that a debate is currently ensuing between two camps of musicological thought where on the one hand “formalist” – those that “advocated semiotic approaches”, are pitted against those that concern themselves with the “methodologies borrowed from hermeneutics and cultural studies.” Since the principle of non-culturally linked hermeneutics that rather signify images between the two emergent scores of *In-Gest* is a primary factor in the structure and ultimately resulting sound-world of the work, it thus approaches the notion of the music being dislocated from cultural or historical reference and rather more aligns with the semiotic, formalist approach.

Further evidence for this aesthetic is provided by the fact that it borrows methods of the incorporation of external sound sources (such as the footsteps in *Walk*) in order to generate new timbres. Thus, the connection created between the acoustic and electroacoustic in *Walk* provided the possibility to include extra-musical elements with musical ones. The sheet therefore introduces the notion that sounds that would

otherwise be disconnected from musical discourse are now connected through their morphing and superimposition with the acoustic materials (Demers, 2010:38). The facility for electroacoustic music to be “mimetic and representational or abstract and obscure” adds impetus to its ability to emancipate itself from established musicological “theories about its roots on culture and history” (Demers, 2010:23).

The aesthetic intention of each sheet by exploring the bounds of timbral limitations of the violin through its extension by the mediation technology, is to concern itself with what is referred to as the “*something to hold on to factor*” (Landy, 2006:30). Since each sheet is interested in providing a gestural affordance – a semantic image of movements in the outside world – the electroacoustic work *In-Gest* is certainly the result of interest by the composer in “discovering meaning in electroacoustic compositions” (Ibid). Consequently, the choice to explore this communication of meaning through the means of the electroacoustic art music aesthetic is supported, since according to the definition as supplied, the medium mixes the disciplines of *musique concrète* and electronic music where the possibilities for exploring musical communication are boundless. Boundless in the sense that harmonic, rhythmic and other micro-structural elements are not limited by the constraints of a language that is defined by its cultural context. Where it fits though, is certainly into the well-established cultural landscape of electroacoustic art music by making reference to established

compositional techniques and methods, a cultural discourse that is clearly concerned with sound and gesture.

Notwithstanding it is the all-important development of a technological heuristic as an integral part of the composition and the aesthetics, which this conveys. As shown in section 1.1 of Chapter 2, the once considered boundaries of art and technology as separate aesthetics have crumbled. The ever-increasing interest in interactive forms and their application in composition, stands testament to this fact. Garnett (2001:32) refers to this in the context of contemporary art music where it

“no longer needs to address the nature of the machine as something outside of us mere mortals, projecting its inhuman capabilities as an aesthetic of the machine; rather, it can now seek to integrate technology into our very selves, and to be intimately controlled by ourselves.”

The example developed here extends this notion, because of the significant capabilities of the system to act as a compositional form of poiesis.

The combination of an analysis and aesthetics consideration of much electroacoustic music is missing in the musicological tradition, which has usually been concerned with the development of tools and techniques that relate to the act of composition. Testament to this is the vast array of

publications available (far too numerous to mention here), that elucidate on the numerous techniques adopted, developed and used in the creation of electroacoustic music, but don't consider the possibilities of poesis and their aesthetic implications (Dahan & Pollet, 2004). Two formal contributions to the literature that plug this gaping hole are the already mentioned Simoni (2006) and Licata (2002) volumes, but glaringly obvious are the scant publications by others that consider both aesthetics and poesis of electroacoustic music.

Whilst the overriding intention of the research presented in this thesis is not to provide a literary equivalent of the analyses present in the published volumes mentioned, the effort may provide the body of electroacoustic music analysis with a fresh example, albeit a subjective one. Furthermore, a note on the investigator's personal perspective, with respect to the autoethnographic method adopted to detail the compositional process and consequently the compositional language and aesthetic, is given space in the concluding chapter in section 7.1.3.1.

CHAPTER 7 - CONCLUSION

Once there was a time when the bringing-forth of the true into the beautiful was called techne. And the poiesis of the fine arts also was called techne.

- Martin Heidegger, 1977

7.1 - In Summation of Processes, Procedures and Models

Throughout this thesis, the exploration of various processes, procedures and models has been given and detailed, with a view to giving insight into a compositional process adopted where the process itself casts the composer as a technologist. Whilst the autoethnographic style of investigation is intrinsically unempirical and subjective, the cognitive processes and compositional analysis presented in previous chapters has been undertaken without prejudice or intentional bias. The overall process of the research undertaken is also given a summary.

7.1.1 - Process Parts: A Summary

The processes can be summarised by dividing them into their various constituents as experienced.

7.1.1.1 - Compositional Ideation

As seen, prior to writing any music, several questions were posed. These acted as agents of creation, where the process of answering them informed the methodology of the creative development. Once formed, the questions enabled a system of poesis to begin, whereby musical forms could underline the answers by producing a work, a process with skills and knowledge acquisition. Following on from the initial questions, perhaps the most critical of these ontological activities, were the musical ideations that occurred as the initial seeds. Firstly by virtue of sketches that resembled inner-eye movements or what has been described as musical Gestalts that invoke gestures.

Next, the sketches were used as notated material to generate electroacoustic transformations to be used as micro-structure data. This in turn informed the electroacoustic score, but also large parts of the acoustic score. How the two scores would be realised simultaneously in performance was the next major consideration.

The idea to develop a mediation technology was spurred from the necessity to find a heuristic that would allow the answering of part of the set of original problem questions posed. This was viewed as an integral part of compositional idea and process, since the technology and musical parts are intricately linked in a form of art/technology syncretism. The syncretism of the two disparate disciplines is a central theme in the original ideation, with the intention to contribute to the research domain.

7.1.1.2 - Compositional Process

Whilst it is certainly true that the development and manifestation of ideas are part of the overall compositional process, there are parts that are noticeably removed from ideation, yet are integral to the composition at hand.

These are annotated below:

- technological tool development;
- programming knowledge acquisition;
- compositional technique research and development: extended techniques, notational systems, audio transformations, procedural and data driven event manipulation and control;
- program troubleshooting and limitation workarounds.

The structure of the course from ideation to compositional process has been outlined previously in section 2.4 of Chapter 6, Figure 50., *Structural Scheme of In-Gest.*

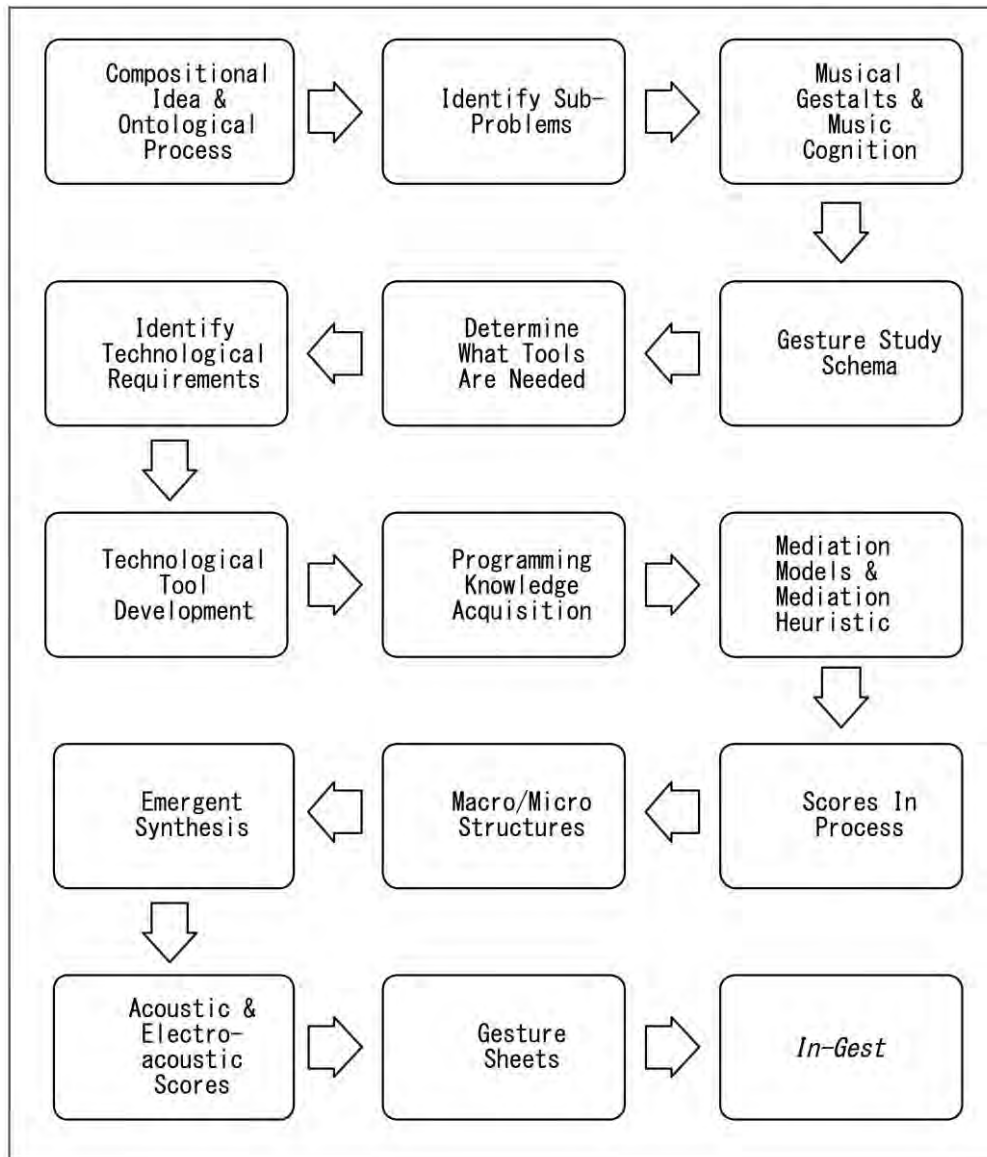
In a final summary, what is required now is a diagram to provide a pictorial outline of the overall compositional process as observed. This is somewhat of an expansion of Figure 1., seen in section 2.3 of Chapter 1. See Figure 51., on page 187. In consideration of this outline, what is apparent are five important overriding components when embarking on a composition that develops and applies a technological heuristic to the problems:

- a study of gesture is critical in developing mediation forms;
- steep knowledge acquisition curve in developing the heuristic must be accepted if a priori does not form part of the individual's referential background;
- programming software that assists in compositional idea forms part of the compositional process and is a form of composition;
- modular (fixed) and procedural (unfixed) macro and micro structures that form emergent synthesis are possible in gesture heuristics and must be considered as a compositional parameter;
- patches, or programming that is directly involved in the music making process should be considered as part of the score.

Value judgments that determine aesthetics are also part of the overall process. Compositional style, which is related to technique either previously acquired or learnt, plays a very important part in the process. The interest for example in manipulating sounds and their properties, with the act of combining them with acoustic instruments, where the two can interact in musical dialogue, is indicative of compositional practice in the electroacoustic art music domain. Additionally, in the sound generation phase, timbral, temporal and harmonic aspects of the sounds generated were implicit in careful consideration of the acoustic score's aesthetic by influencing the pitches, harmonic structures and rhythms.

The choice of software and systems also had an effect on the overall aesthetic by virtue of the separate systems (audio and video) influencing the structure of process, such as the ability for the system to use specific live events to create sonic events. Similarly, where the system creates the possibility of indeterminate values from the tracking data, a somewhat aleatoric aesthetic presents itself. This is a property, which is viewed in summation, as a core part of the indeterminate primary tertiary macro structure of the work, where the final narrative is determined from performance to performance.

Figure 51. – Compositional Process Summary



7.1.2 - Research Summary

What follows is a summation of the research process and overall outcomes. Owing to the nature of this thesis being partner to a portfolio, and in particular the work *In-Gest*, the research was undertaken as an appraisal

of the process to realise *In-Gest* and to also provide insight into the musicological findings resulting from the literature review.

The literature review undertook to support the methodologies and rationales proposed. The review is summarised by topic below:

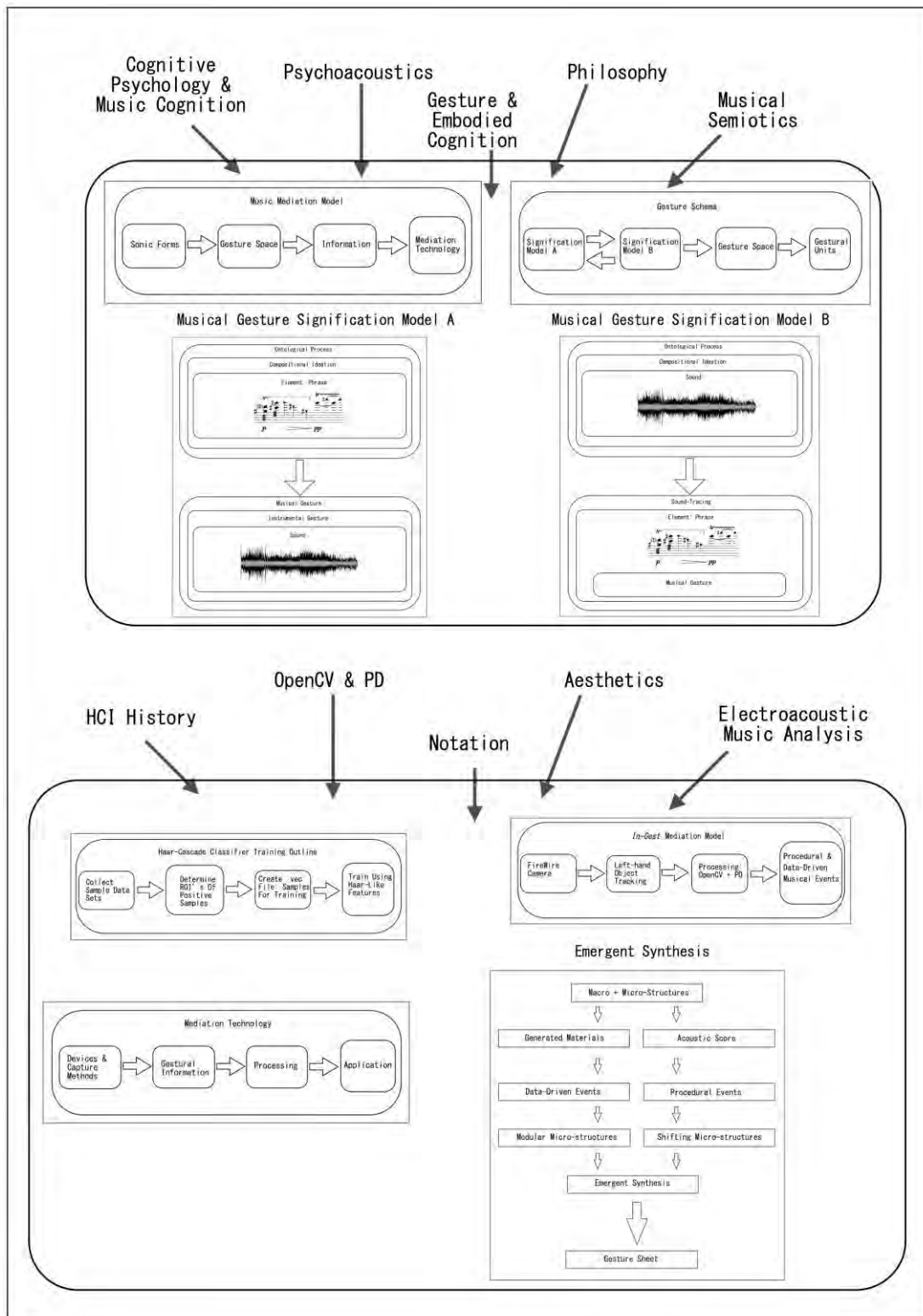
- Cognitive Psychology, particularly Music Cognition
- Psychoacoustics
- Musical Gesture
- Embodied Cognition
- Philosophy, principally that of Phenomenology
- Musical Semiotics
- Examples and History of HCI in Musical Domain
- Technology, particularly OpenCV and PD
- Notation, particularly extended and unconventional forms
- Aesthetics relating to Electroacoustic Art Music
- Electroacoustic Music Analysis

The extensive survey allowed for the proposed models and methodologies to be supported. The diagram, in Figure 52., on page 190, illustrates how the literature survey encompasses the models proposed. Clear from this, is that each of the broader topics fed the establishment of the various methodologies and creation of models that resembled sections of the

compositional process. Thus the research, as stated above, acted as a critical component of the composition itself.

Whilst it is certainly very difficult to pin-point published literature with the specific topic of 'research as a part of compositional process', there is certainly many examples of the application of it without actually referring to it as such. For example, Laske (1999:155) defines the compositional process as "including the learning process" comprising of various steps, the first of which is "learning the available software and the peculiar control variables of its interface(s)", and so on. What this intimates is, that if vocational studies are often critical components of the compositional process, then it stands to reason that higher-level studies such as complex musicological subjects may also form part of the process. This was certainly the intention in the research undertaken as presented in this thesis.

Figure 52. - Literature Review Topics and Proposed Models



7.1.3 - Composer-Technologist Transpired

In previous chapters, particularly 2, 5 and 6, the autoethnographic investigation of composition as a process involving technology revealed three over-arching discoveries:

- the composer must initially embark on an extensive literature survey regarding the possible applications of technology to the idea;
- the composer must tackle technical issues through a process of trial and error;
- finally, once the above has been achieved, the composer is in a position to apply the skills and knowledge gained to solve the problems.

As defined, technologists are experts in their field. They do not necessarily have cross-discipline skills and therefore, in the purest sense, the definition does not include the possibility of individual experts possessing multi-disciplinary skills, particularly where they are vastly disparate, such as music and science. However, where it occurs that a multi-disciplinary skill-set transfers knowledge from one to the other by a hermeneutic cycle (as demonstrated in chapters 4 through 6), the portion of the individual's skill set that was developed in tandem to the existing skill-set, places one within the scope of the technologist definition, since it now obtains an additional term, i.e. *composer* is

added. The technological expertise gained is concomitant and reliant on the existing skill-set, since as shown, they are developed in tandem. This provides sufficient impetus to the claim that the composer becomes a technologist when adopting self-learnt expertise and skills in solving compositional problems that require the development of technological heuristics. It should also be noted, that since technology is a constantly evolving phenomenon and that the system developed as presented in this work is intended to undergo future enhancements (see section 7.1.4 below), the spirit of technological evolution and the composer-technologist's work are also by default in a state of constant development and refinement. The process as described and detailed in this thesis has certainly shown this to be the case.

7.1.3.1 - Compositional Process as Technologist: a Note from Personal Perspective

The descriptions and summations presented thus far have been by and large, analytical but focusing on the contextual work relating to the research. It is intended that in this section a closer look be given to how working as a technologist affected the overall compositional process of the investigator and critically the effect this has had on the musical outcome of a composer whose experience of this mode of process was entirely new.

The author's prior compositional background and output was almost entirely principally acoustic. Preceding to this investigation, there are only two

other works that either incorporate electronics with an acoustic instrument or are themselves exclusively electronic. A work that falls into the former category is Warrington (1977) *A Dialectic Study* for piano and tape (1999 & 2000)⁶³. Composed in 1999 with substantial edits and additions in 2000, this work involves no interaction, control or manipulation of the tape part by the performer and is of a purely dialogue/accompaniment format between the two parts.

The other is a portfolio of entirely electronically produced compositions accompanying a research article investigating computer game music - see Warrington (2005b). Out of these two examples, the latter body of work would best fit into the 'collaborating with technologists' paradigm. The technologists in this case were the team working on a project called the Virtual Learning Space (VLS) project. For a detailed overview of this project, the reader is directed to Baxter, Seagram & Amory (2006).

The works created for this project, were entirely reliant on third party input (computer user) in order to be realised. This is because of their interactive nature and their dependence on the position of the game player's avatar. This non-linear interactive nature of the narrative produces some unique compositional problems that are quite specific to gaming audio or other non-linear mediums that are directly influenced by third-party manipulation. It was a unique experiencing writing for this

⁶³ A complete score of this work is available on request through the contact form at <http://www.mileswarrington.com>.

medium, and as attested in Warrington (2005a), the process informed the output style of the compositions significantly. For example, it is noted that where the composer wished to achieve particular results, the overriding and mitigating factors that determined this, were entirely the nature of the technology used, where in Warrington (2005a:42) and (2005b:47), it is suggested that a theoretical game audio engine be created in order to realise the goal of a completely successful interactive composition.

Subjectively speaking this was challenging at the time, since the software responsible for the game audio was being developed in parallel and almost simultaneously with the music for the project. This meant that there were often situations where completely different musical outcomes would be achieved instead of what the composer had in mind once the technology specific to a scene or area in the game had been completed. Some sections were completely re-composed and indeed, to this end, a complete sequence of audio based in an "imaginary sequence" was added to the body of work to demonstrate this (Warrington, 2005b:13,45). The interactive nature of the compositional work undertaken in the VLS project gave the researcher some experience in the medium of interactive works. However, the fact still remains that where a performer is involved as the principal controller, and whom is also responsible for musical output, a completely new dynamic occurs where the composer must be simultaneously aware of each side and their influence on one another. This was certainly the case in the

compositional undertaking to create *In-Gest*, and is one of the critical differences between this and the older works, provided for by the use of the technology developed in the compositional process.

The tables that follow, detail in summary form the electronic works composed prior to *In-Gest* and its composer-technologist experiment. Also given are compositional techniques, which were contingent on the technology used at the time. An observation on that relationship is given below each table, and these are compared with the observations that follow-on from Table 7.

Table 5. – Electronic Works Prior to *In-Gest: A Dialectic Study* (1999 & 2000)

Work	Technique(s) Employed	Technology Developed/ Implemented
<i>A Dialectic Study</i> for piano and tape (1999 & 2000).	<ul style="list-style-type: none"> • Sampling: recording, construction and playback; • Layering; • Articulations between acoustic and linear tape-track. 	<ul style="list-style-type: none"> • Compact Disc or computer with audio playback functionality.

Observations:

In this, the first work written for the electroacoustic medium by the researcher, the techniques present were limited only by the experience in

working within the medium at the time. Evident from this is the fact that no time stretching, spectral layering, partial extraction nor live synthesis techniques were applied to or used in the samples present on the CD as the accompaniment. A few short sections of layered samples processed using a granular synthesis application⁶⁴ occur in short sections of the piece. The overall harmonic language of the work also leans to a tonally-centered dialogue between the acoustic instrument and tape part. The samples themselves are that of an African bamboo flute, recorded live and used in strict time with the written acoustic piano parts on the CD. It is clear that at the time the technology used in the piece had not directly influenced the composer's musical aesthetic, since the overall language of this work and that of other purely acoustic works⁶⁵ from the same period are similar.

Table 6. – Electronic Works Prior to *In-Gest: gammaKhozi: Burning Ground & VLS Game* (2005)

Work	Technique(s) Employed	Technology Developed/ Implemented
<i>gammaKhozi: Burning Ground</i> compositions for the VLS Game (2005).	<ul style="list-style-type: none"> • Recording of ambiances; • MIDI-based sample triggering; • Multi-track envelope 	<ul style="list-style-type: none"> • VLS Game audio engine.

⁶⁴ At the time this was Granulab 1.0. More details are available at <https://www.abc.se/~re/GranuLab/Granny.html> (accessed 2016, April 25).

⁶⁵ For details of the works, an anthology of compositions for the composer's portfolio submitted for degree purposes in 2000, is available in hard copy from the Eleanor Bonner Music Library, University of Kwa Zulu-Natal, Durban.

Work	Technique(s) Employed	Technology Developed/ Implemented
	control; <ul style="list-style-type: none"> • Basic spatialisation techniques. 	

Observations:

For the first time in the author's compositional practice, the use of a technology to realise a compositional goal had a direct influence on both craft and aesthetic. Working with a game engine that allowed a non-linear pathway through the virtual narrative, created a situation where underscoring sounds to a scene had to merge seamlessly with the underscoring of the previous scene and the following scene. Subsequently, with each new scene being added the same technique was utilised to maintain the seamless flow. Whilst this was one desirable outcome, it did not allow for a complete realisation of the ultimate goal for composition at the time. This was due to the fact that individual compositional segments could not be manipulated in real-time by the engine due to its inherent limitations. This meant that the composer had to theorise an alternative version of the music in an "Imaginary Sequence"⁶⁶. This resulted in an entirely aleatoric musical aesthetic, which was directly influenced by the theorised technological solution.

⁶⁶ See Warrington (2005b:13, 45).

In-Gest is itself an interactive work, but the compositional outcome due to its process was never contingent on the final interactive result i.e. during or after performance. At the very outset of the experiment it was always apparent that a 'hermeneutic' process would unfold during the creation of the work. This means that the construction of the heuristic in parallel with the music has directly informed the aesthetic context of the work and therefore also the composer's language with respect to it. It is consequently evident that the transpiration of the composer-technologist phenomenon has informed the compositional aesthetic of *In-Gest*. This has meant, very critically, a distinct departure by the investigator from previous compositional aesthetics associated with the two previous works that incorporate electronics and indeed those that are purely acoustic in nature. This is wholly evident from several musical factors that have been identified in the musical output from it. To contextualise these within this autoethnographic perspective, the tables below correlate in summary form, a salient example from each of the gesture sheets that indicate these language and aesthetic changes, with the compositional techniques as identified in section 6.2.3, and the 'new' process of composition with those already identified in section 7.1.1.2 of this chapter. As above, a personal observation is given below each table. It is intended that the observations detailed will provide a link between the emerged composer-technologist practice and the actual act of composition, particularly how that has changed for the researcher. This is presented, again, in order of analysis as occurs in the thesis.

Table 7. – Compositional Process: The Composer–Technologist of *In–Gest: Walk*

<i>In–Gest</i> Motif/Section/Phrase	Technique(s) Employed	Technology Developed/ Implemented
<i>Walk</i> : From start of [detection 1D] until end of [detection 2P].	<ul style="list-style-type: none"> • Overlaying; • Relating differing sound materials; • Articulations between acoustic and procedurally created sounds; • Overall gestural links. 	<ul style="list-style-type: none"> • Timed detection events; • Audio partial extraction techniques; • Audio feature isolation.

Observations:

The use of untransformed acousmatic material here with purely acoustic material, was a first–time occurrence for the composer. The technology permitted two disparate sounds to influence each other through timing and therefore ultimately establish an unusual sound–rhythm⁶⁷. Exploring the well–established technique of audio feature extraction has preempted the use of other spectralism techniques in this and future compositions that follow, a hitherto unused compositional technique by the composer.

⁶⁷ ‘Sound–rhythm’ is used here instead of the conventional ‘harmonic rhythm’ term, but intimates the same meaning implied by the latter, only ‘harmonic’ is a less suitable term in the context of the work’s sound–world.

Table 8. – Compositional Process: The Composer–Technologist of *In–Gest: Up*

<i>In–Gest</i> Motif/Section/Phrase	Technique(s) Employed	Technology Developed/ Implemented
<i>Up</i> : [detection 2P] and [detection 3D].	<ul style="list-style-type: none"> • Mimesis; • Overall gestural links. 	<ul style="list-style-type: none"> • Timed detection events; • Spectral analysis - sonogram use; • Auditory illusions.

Observations:

The use of sonograms to identify and generate notation, out of which sound structures could be created was an entirely new technique adopted for the researcher. Whilst this is common–place amongst regular practitioners of the electroacoustic music art form, the adoption of this technique has greatly helped to provide the researcher with better insight into the micro–structuring of sound events that themselves lead to larger structures. Additionally, the application of an ‘auditory illusion’ in the form of the Shepard–Risset Glissando present in this sheet, was another yet unused compositional technique by the investigator. Critically, all of these techniques have occurred as a direct result of applying the heuristic developed as part of the compositional process.

Table 9. – Compositional Process: The Composer–Technologist of *In–Gest: Down*

<i>In–Gest</i> Motif/Section/Phrase	Technique(s) Employed	Technology Developed/ Implemented
<i>Down</i> : All detection points.	<ul style="list-style-type: none"> • Articulations between acoustic and procedurally created sounds; • Assimilation; • Overall gestural linking. 	<ul style="list-style-type: none"> • Timed detection events; • Time–stretching. • Real–time randomisation.

Observations:

The use of an extended technique on the violin to produce a ‘subharmonic’ had never before been undertaken by the scholar. Indeed, this specific technique is the most ‘radical’ present in the work in terms of the acoustic score. Although this is the case, the most striking unfamiliar procedure in terms of previous compositional process in this regard, is the use of a recording of such a technique together with a randomisation engine, to produce a very specific effect. It would be very difficult to create a similar effect in a purely acoustic setting. The narrative in *Down* demanded that it sound as if the acoustic part is essentially ‘tripping over’ the electroacoustic part as it rolls down an imaginary hill. The essential part is the tracking technology allowing the timing between events to create the perception of random ‘tripping’.

Table 10. – Compositional Process: The Composer–Technologist of *In–Gest: Waltz*

<i>In–Gest</i> Motif/Section/Phrase	Technique(s) Employed	Technology Developed/ Implemented
<i>Waltz</i> : All procedural Detection points: [2, 3, 7, 8, 11 & 13P].	<ul style="list-style-type: none"> • Articulations between acoustic and procedurally created sounds; • Assimilation; • Overlaying; • Overall gestural linking. 	<ul style="list-style-type: none"> • Timed detection events; • Real–time audio envelope manipulation.

Observations:

This was the first time in a composition by the investigator where pre-composed and recorded material has its fixed audio data content changed in real-time based on a performer’s input. The unique nature of the score in this sheet, allows for the performer to control the timing of what is termed ‘motivic interjections’. Compositionally, this has opened up the possibility of thinking of sound as disjointed objects in time, whose inherent structure is wholly dependent on the temporal localities of the neighbouring sounds.

Table 11. – Compositional Process: The Composer–Technologist of *In–Gest: Float*

<i>In–Gest</i> Motif/Section/Phrase	Technique(s) Employed	Technology Developed/ Implemented
<i>Float</i> : Both detection points.	<ul style="list-style-type: none"> • Outside time 	<ul style="list-style-type: none"> • Timed detection

<i>In-Gest</i> Motif/Section/Phrase	Technique (s) Employed	Technology Developed/ Implemented
	structures; <ul style="list-style-type: none"> • Overlaying; • Relating different sound materials. • Layering; • General gestural linking. 	events; <ul style="list-style-type: none"> • Time-stretching.

Observations:

The very idea of previously exploring extreme time-stretching was altogether foreign to the author's technique. With the determination of a performer's timing between events effecting the relationship of the duration of the data-driven events to that of the acoustic score, the technology provided a unique opportunity to explore time both as phrase length and as the sustain of notes. This has afforded the researcher with the tools and insight to explore time envelopes of sound in more detail in future compositions, a clear and formative property related to the aesthetic transformation under discussion.

Table 12. – Compositional Process: The Composer-Technologist of *In-Gest: Slow*

<i>In-Gest</i> Motif/Section/Phrase	Technique (s) Employed	Technology Developed/ Implemented
<i>Slow</i> : [detection 3P]	<ul style="list-style-type: none"> • Mimesis; • Overlaying; 	<ul style="list-style-type: none"> • Timed detection events;

<i>In-Gest</i> Motif/Section/Phrase	Technique(s) Employed	Technology Developed/ Implemented
	<ul style="list-style-type: none"> • Articulations between acoustic and procedurally created sounds; • Overall gestural linking. 	<ul style="list-style-type: none"> • Real-time control of granular synthesis.

Observations:

This section of the work contains the implementation of the first occurrence of live control of a synthesis technique in the composer's oeuvre to date. This is a direct result of the heuristic developed and indicates a novel application of it in terms of compositional technique in the author's contextual output. In reflection, there is no doubt that the technology developed is directly responsible for compositionally thinking of linking musical gesture in this way.

Table 13. – Compositional Process: The Composer–Technologist of *In-Gest: Fast*

<i>In-Gest</i> Motif/Section/Phrase	Technique(s) Employed	Technology Developed/ Implemented
<i>Fast</i> : Detection points: [2, 3, 5, 6, 8, 9 & 10P].	<ul style="list-style-type: none"> • Mimesis; • Layering; • Relating differing sound materials; • Articulations between acoustic and 	<ul style="list-style-type: none"> • Timed detection events; • Artificial 'variable delay' application in real-time.

<i>In-Gest</i> Motif/Section/Phrase	Technique(s) Employed	Technology Developed/ Implemented
	procedurally created sounds; <ul style="list-style-type: none"> • Overall gestural linking. 	

Observations:

As with the auditory illusion present in *Up*, the artificial Doppler effect created in this section of the work is just another example of the technology informing compositional practice that makes use of such devices. It was intended that there is a mimetic link with the indication for the performer to accelerate the bow without changing dynamic. Through the technology's ability to allow this mimesis with the acoustic score in the electroacoustic score, the composer was able to explore that relationship for the first time. This is a particularly exciting development, since wherever it is possible to achieve similar effects on other acoustic instruments, it opens up the possibility for exploring analogous compositional practice in future electroacoustic works with live instruments.

The tables above have endeavored to make potent the link between the emergent composer-technologist in the context of the autoethnographic compositional process detailed in this thesis, and the author's respective change in compositional technique, language, aesthetic and ultimately practice.

7.1.3.2 - Brief Summary of Observations from Tables 5., through 13.

Glancing a comparative eye over the observations from above, it seems obvious that where the composer was directly involved in developing or utilising technology to control compositional outcome and therefore control process, the aesthetic constitutes some dramatic changes. We see a shift from a generally conservative harmonic language and sound design in the *Dialectic Study* where minimal technology was consumed and none developed, to a more adventurous aesthetic in the music for the *gammaKhozi* game. In this latter example, the technology was consumed in a parallel fashion as a product of innovation and development by the VLS team, leading to some interesting compositional outcomes as detailed. The most pertinent difference in outcome here was the theorisation of a technological solution by the composer to realise the ideal compositional objective, itself leading to an entirely different musical aesthetic in the form of aleatoric music for the computer game. Lastly, once the researcher had reached the level of developing technological heuristics as part of the compositional process itself (as detailed in the investigation of the creation of *In-Gest*), we see how assuming the technologist role had drastically changed the investigator's compositional language and altogether informed compositional practice.

The results here echo the earlier sentiments proposed, where the composer's perception in the act of reception is intentionally objective, however it

may be 'guided' with techniques (referring to technology), that create particular aesthetics.

7.1.4 - Future Work

As discussed, it is intended that the system developed is continually refined to produce more reliable object identification and tracking. This will involve several undertakings:

- solve the PD crash problem to allow the use of the higher specification camera. If not, port to MAX/MSP;
- use a cloud-based computing array or other such powerful multi-processor system to complete the training of the large 160X160 vec file;
- enhance the existing OpenCV binding for haarcascade classifiers for PD or create an entirely new one that doesn't drop the object identifier once tracking is determined;
- develop a fully packaged portable version of the system that allows composers to easily setup and apply it in a wide array of interactive music applications and package the system to allow multiple performance possibilities of *In-Gest*;
- embark on a fully professional realisation of *In-Gest*;

- work with technologists in the robotics and engineering fields to refine the current system as well as possibly developing alternative methods of musician gesture tracking and recognition systems.

Besides the technical improvements that go hand-in-hand with the intended future work, it is anticipated that a reference book on the application of Computer Vision as a compositional tool be written. This will have particular focus on the combined use of OpenCV and object-orientated programs such as PD and MAX/MSP, which are ubiquitous, well tested and known within the music community.

7.2 – Concluding Remark

At the end, the research presented here has achieved three goals:

- cast the composer as a technologist through a process of compositional problem solving;
- point-up that original music composition is an important form of academic research;
- and position the research at the divide of the debate ensuing between the art and science disciplines.

Technology has been shown to be a form of poiesis both when used as a heuristic to compositional problems and when applied in ways that make it

indispensable in the musical work. This outcome and those annotated above, positions the work at the forefront of innovation, particularly the emerging field of gesture-related studies that incorporate motion capture, object tracking and recognition. Aesthetically, it provides the genre with a work that attempts to reveal the importance of a composer's awareness of the ontological processes that present musical forms as movement or gestures that can be given meaning within the framework of sound. Finally, it provides scope for the application of technology in composition as a viable and necessary pool of philosophy.

APPENDICES

APPENDIX I - *In-Gest* Full Score

APPENDIX II - Main Audio Patches

APPENDIX III - DVD DISC

APPENDIX IV - Hand Model Contributors

APPENDIX I - *In-Gest* Full Score

Notation Guidelines And Symbols For *In-Gest*

Violin Score:



Play molto accelerando or decelerando as indicated.



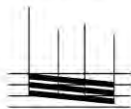
Bi-Directional Glissandi.



Play in one bow-stroke with acceleration in the stroke and at a constant dynamic, i.e. no crescendo.



Left-hand pizzicato on ad-lib. notes within rhythm notated.



Head-less rhythm: Holding the bow vertically, dab the approximate pitches indicated with the screw of the bow.



Stemless notes: Duration of notes is ad libitum.



Special block notehead with stem: Execute extreme bowing pressure on partial node as indicated to produce "sub-harmonic". Duration of sound is determined by stem i.e. quaver in this case. If there is no stem above the block notehead, the duration is ad libitum.



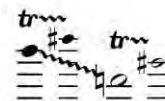
"Sub-harmonic" or Anomalous Low Frequency (ALF). An approximation of the pitch produced by the extreme pressure bowing on the G string at indicated nodes.



Inverted triangle note head: play flautando.



Trill with small notehead.



Trilled glissandi: always with sul. pont., and augmented fourth intervals in the trills.



These indications occur above notes/chords where HCI detection takes place. The Roman numeral indicates string i.e. I-IV represents E-G respectively; the number indicating finger. The note/chord must be fingered as indicated.



As above, only this indicates the fingering and string required for a chord - the top set being the top note and bottom the lowest note in chord.



Alternating pizz. and arco motifs. The 'x' above the note indicates play pizzicato. This is followed by an upward bow marked 'v'.



Raise pitch indicated by 1/4-tone.



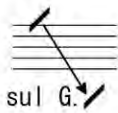
Raise pitch indicated by 3/4-tone.



Gradually change from sul pont. technique to nat.



Gradually merge into next event. Occurs especially during ad lib. sections.



Execute a slide on the string indicated between the two approximate pitches as shown.



Execute an extremely fast bowed-tremolo.



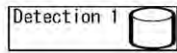
Gaps in staves: time may be taken (ad lib.) between the sections.

Note on accidentals: Because there are no barlines nor key signatures in this work, the standard rule applies where each new group serves as the point where an accidental must appear before the note even if the same note is used from the preceding group.

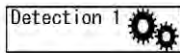
PERFORMANCE NOTES:

The piece is constructed of "Gesture Sheets". Each sheet may be played in any order chosen but may not be repeated during the same performance. There are no time signatures nor barlines in the work, as all the musical material is meant to be free from temporal constraints and beats unless specific instruction is given, for example: ritardandos, accelerandos and tempo indications where indicated. The tempo markings are mere guidelines and are not meant to be adhered to strictly, unless indicated to the contrary.

Electroacoustic Score:



Approx. point of detection of sound-producing gesture by camera system where the system uses data-driven processes to control the resulting sound, for e.g. specific fingering triggers a sound streamed off the hard-disc drive. The detections are numbered according to relative acoustic-score positions and to provide reference to the detailed score analysis.

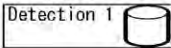


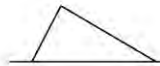
Approx. point of detection of sound-producing gesture by camera system, where the system uses procedural processes to control/morph or create the resulting sound, for e.g. timing between events, specific (X,Y) positions of fingers or relative size (Z) of detected object to Z axis of 2 dimensional plane.



When this symbol is indicated, it delineates procedurally processed audio where the start and end limits of the audio in question would otherwise be difficult to determine (such as sonic events inside a box). The direction of the symbol indicates either start or end. In this case, "Start".



"End" version of symbol above. "end" points usually controlled by  if indicated.



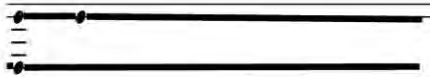
Pre-determined attack & decay envelope triggered by specific sound-producing gesture using data-driven method. The envelope is existing in the data-driven audio stream.



Attack & decay envelope whose shape is procedurally determined by the timing between detection events. This is then applied in real-time to transformed sonic events.



Stem-less notes: Lack of stems indicate the possibility for indeterminate timing of sonic event. The overall duration is however meant to be an accurate indication.



Stem-less notes with thick lines: As above indicates indeterminate timing, but duration is approximated by the line.



Right-arrow indicates possibility of merging sound events.



Wavy line: This indicates a gradual and un-measured sliding time-stretch that is either processed live (in box) or pre-processed. Pitch is only affected if the original sample contains sliding pitch data. Direction indicates whether it is an increase or decrease.



Procedural partial graphic. Partial is extracted and played back by system from pre-recorded violin sound. Variables determined by previous timing between detection events.



Double-Arrow. Indicates an inside-time accelerando between chords/notes that is pre-processed. If in box, it is processed procedurally.



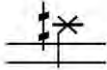
Slashed note-head. No absolute pitch is intended and the attack is imperceptible.

(C)

Bracketed note name. Indicates the approximate root pitch of the sonic event.



Pitch of sound rises and lowers slightly but with a general trend towards sharpening of the note/chord.



Approximate pitch as indicated is raised by 1/4-tone.



Approximate pitch as indicated is raised by 3/4-tone.

13,264kHz



Notated 'sonogram' information. In sections where this appears, the frequency in Hertz of the sound is indicated along with its relative amplitude represented by the thickness of the line.



Approximate time duration of data-driven audio event indicated. When boxed, the beginning and end indicated with audio-wave symbol as described above does not delineate the exact start and end points in the score, but is used merely as a symbol of the event.

INSTRUCTIONS FOR CONTROLLER/SET-UP:

COMPUTERS:

Knowledge of Pure Data (PD) is required. Two computers connected via LAN are needed, running the same version of PD. Place a copy of the VIDEO patch on the one (VIDEO STATION) and AUDIO patch on the other (AUDIO STATION).
Run the patches: Enable video rendering + load classifier file and check connections. Enable each of the AUDIO patches as required in performance from MAIN patch window toggles. Not enabling will result in ZERO data transmission from the VIDEO STATION.

AUDIO:

Prepare a copy of each gesture sheet and enable the 'Panic Patch' in a separate window from the folder of parent patches.
A contact-microphone or pick-up is ideally required on the violin, mixed with the output from the AUDIO station.
A reverb effects processor must be added prior to stereo out, preferably a convolution unit. Amount of reverb is at the performer/controller discretion.
A speaker system with a subwoofer or support for low frequency output is essential.

VIDEO:

A fairly bright incandescent light should be positioned directly over the performer's head to allow for best lighting conditions. Avoid the casting of shadows onto the fingerboard.
Place the camera using an adaptable stand as parallel to the finger board as possible to achieve the best frontal position image capture.

Walk

GESTURE: "WALK"

Violin: $\text{♩} = 72$ IV 1 IV 1
cal legno
mf espress. *f* *mf* *pizz.* *p* *mf*

HC1: Detection 1 Detection 2

ELEMENTS FOR RANDOM FOOTSTEP AUDIO GENERATOR

FOOT 1	FOOT 2	FOOT 3	FOOT 4
Heel	Heel	Heel	Heel
Toe	Toe	Toe	Toe

Violin: *nat. arco* IV 1 IV 1 *mp* *p* *pp*
Detection 3 Detection 4
mp legato

HC1: *mp legato*

Violin: *nat.* IV 1 *mp molto legato* *ppp* *mf espress.*
Tempo ad lib.
IV (sempre sul G)
x (C) x
x (C) x

HC1: *mf* *mf*

Violin: *mp legato* *ppp*

HC1: *mp legato* *ppp*

Up

$\text{♩} = 58$ or slower but not less than $\text{♩} = 44$

free flowing

GESTURE: "UP"

Violin: *sul pont.*, *sempre non vib.*, *mf*, *f*, *mp*

HCI

Detection 1

16.634kHz, 14.728kHz, 13.781kHz, 13.409kHz, 12.833kHz, 11.886kHz, 11.563kHz, 10.938kHz, 9.991kHz, 9.716kHz, 9.043kHz, 8.570kHz, 8.333kHz, 7.149kHz, 6.675kHz, 6.492kHz, 5.254kHz, 4.780kHz, 4.645kHz, 3.359kHz, 2.685kHz, 1.464kHz, 1.937kHz

Violin: *mp*, *f*, *mp*

Detection 2

18.001kHz, 15.633kHz, 13.264kHz, 10.938kHz, 8.570kHz, 6.675kHz, 4.306kHz, 1.937kHz

Violin: *f*, *f*, *mf*, *f*

Violin: *nat.*, *pp*, *ff*

HCI

Detection 3 (Shepard-Risset Glissando)

ppp, pp, p, mf, f, ff, ppp

Waltz

GESTURE: "WALTZ"

♩ = 66

Violin I: *f*, *nat.*, *sul G*, *sempre marcato unless otherwise indicated*, *p*, *f*

Violin II: *p*, *f*, *p*, *f*, *mf*, *mf*, *f*, *sul D.*

Violin III: *nat.*, *(sempre marcato)*, *mf*, *f*, *mf*, *f*, *più f*

Violin IV: *mf*, *mp f*, *mp*, *pp*, *f*, *ff*, *mf*, *f*

Violin V: *nat.*, *(sempre marcato)*, *mf*, *pp*, *f*, *ff*

HCI (Violin I): Detection 1, 2, 3

HCI (Violin II): Detection 4

HCI (Violin III): Detection 5, 6, 7, 8, 9

HCI (Violin IV): Detection 10, 11

HCI (Violin V): Detection 12, 13, 14

Chords: (C), (G#), (CHORD), (A#)

Violin V: *poco a poco martellato*, *nat.*, *(sempre marcato)*, *mf*, *più f*, *ff*

Float

GESTURE: "FLOAT"

Tempo ad lib.
sempre con sord.

Violin
ppp molto legato
poco cresc. pp
p *pp*

HCI
 Detection 1

====

Vln.
p *dim.*
pppp zephyr-like
pp molto legato

HCI
 Detection 2

====

Vln.
mp *pp*
ppp molto legato
pp *p* *p*

HCI

====

Vln.
pp molto legato *p*
p

HCI

====

Vln.
ppp

HCI

Fast

GESTURE: "FAST"

The score is divided into three systems, each with a Violin (Vln.) and HCl part. The Violin part features a complex melodic line with various fingerings (IV 1, I 4, III 4, III 4, IV 2) and dynamics (f, ff, sempre detache). The HCl part provides harmonic support with chords and textures. Detection points 1-4 are in the first system, 5-8 in the second, and 9-11 in the third. The second system includes markings for 'arco', 'nat. pizz.', and 'simile'. The tempo is marked as 'Fast' with a metronome marking of quarter note = 80.

Violin: $\text{♩} = 80$, *f*, *ff* *sempre detache*

HCl

Violin: *arco*, *nat. pizz.*, *simile*

HCl

Violin: *arco*

HCl

Detection 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11

APPENDIX II - Audio Parent Patches

Figure 53. - *Walk* Main Audio Patch

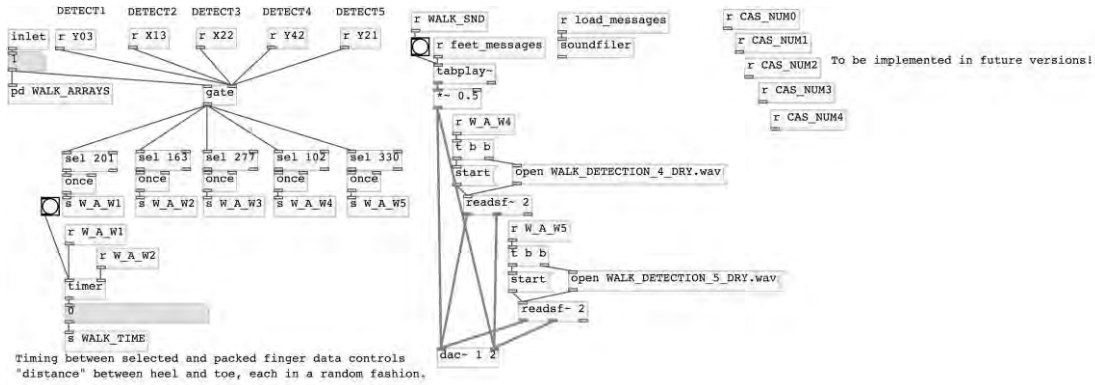


Figure 54. - *Up* Main Audio Patch

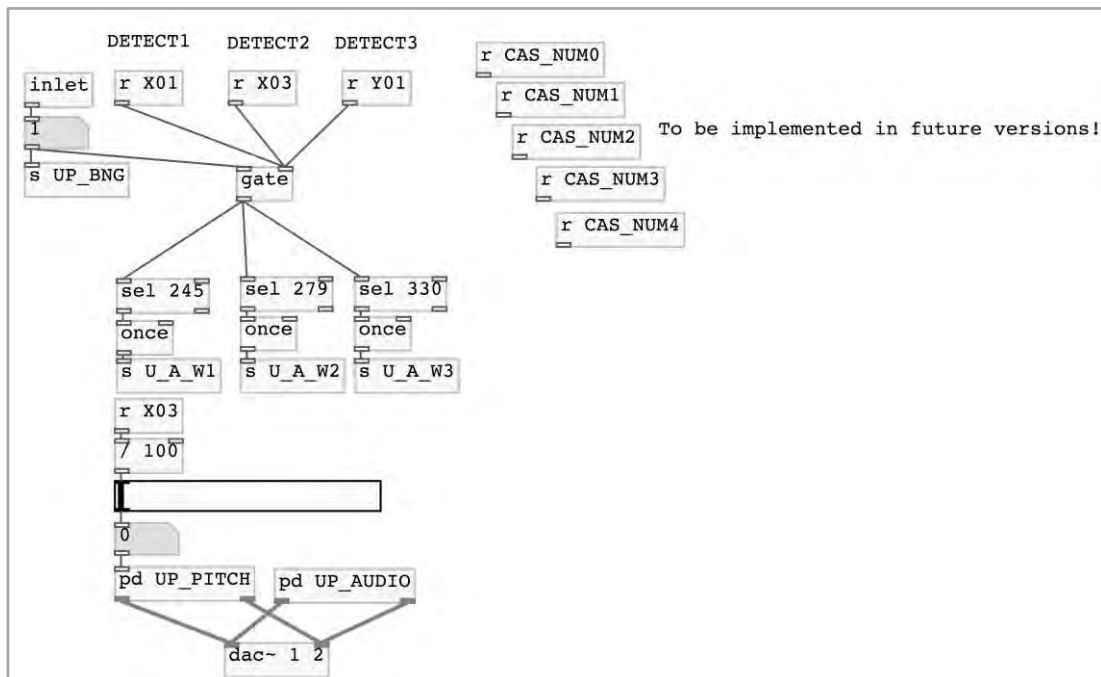


Figure 55. - *Down* Main Audio Patch

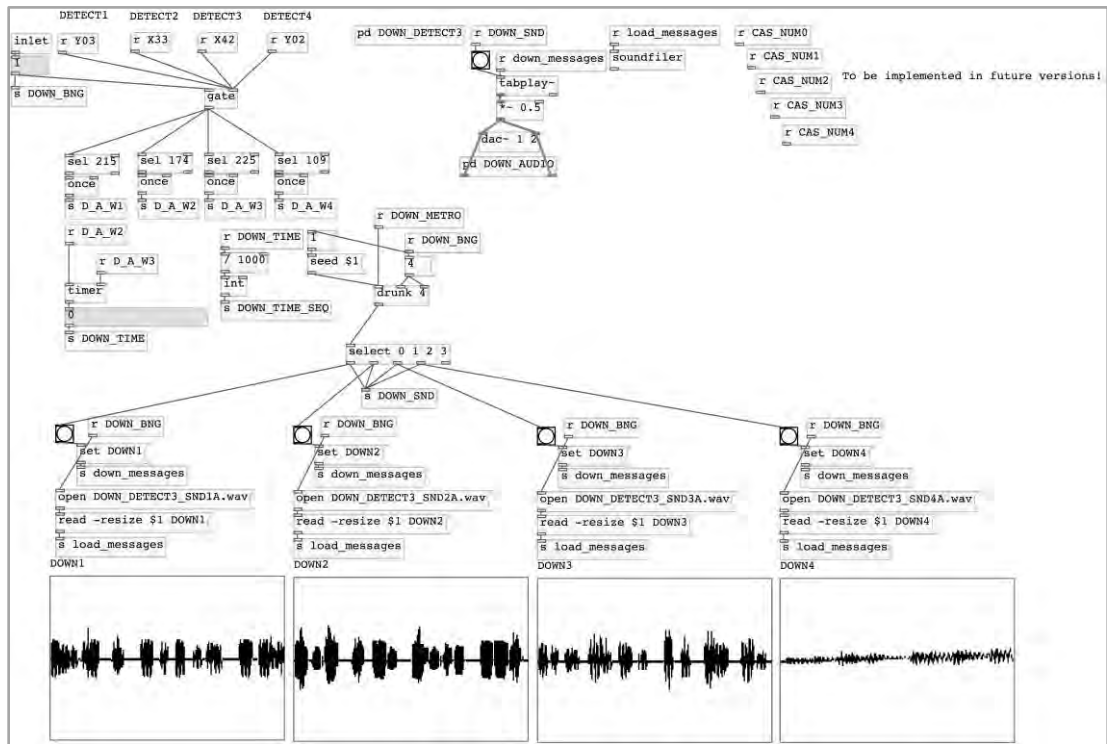


Figure 56. - *Waltz* Main Audio Patch

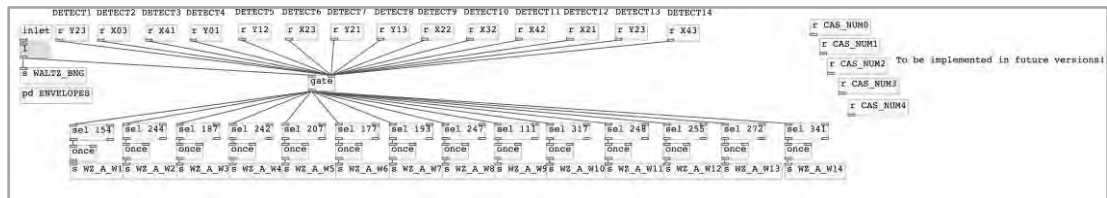


Figure 57. - *Float* Main Audio Patch

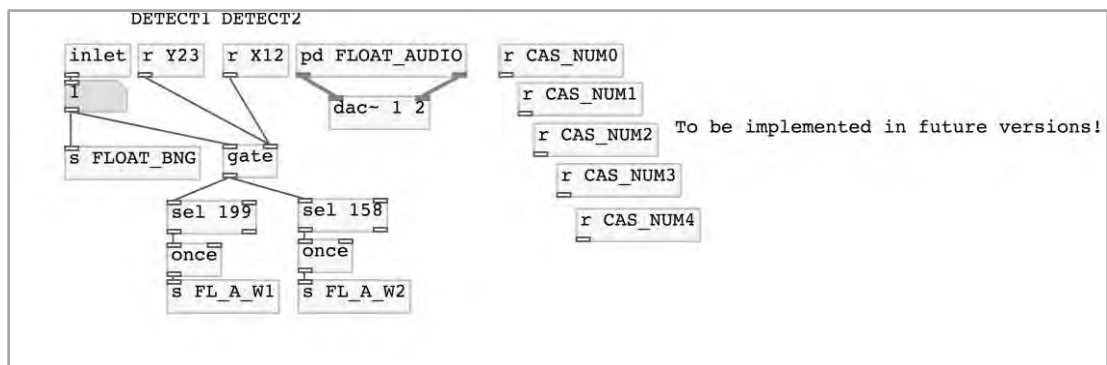


Figure 58. – *Slow* Main Audio Patch

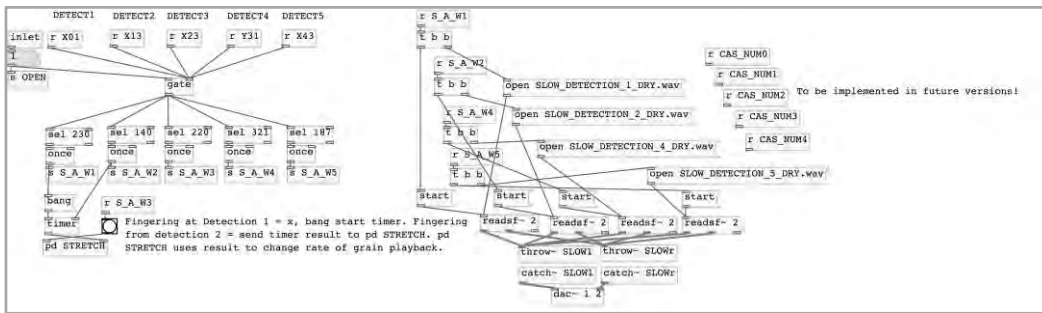
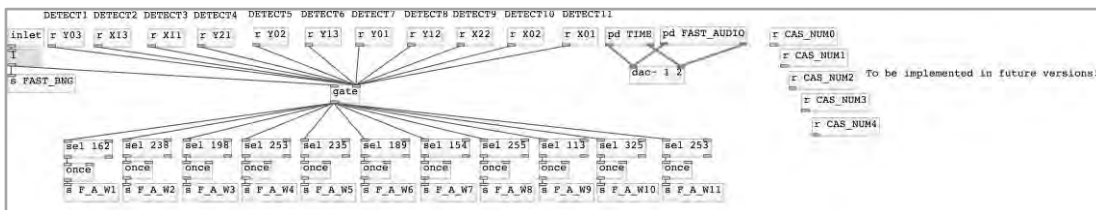


Figure 59. – *Fast* Main Audio Patch



APPENDIX III - DVD ROM

DVD DISC CONTENTS AND FOLDER LAYOUT

Patches:

D:/IN_GEST

Demo Video Clips:

D:/DEMO_CLIPS

Readme

Slow

Up

Walk

Audio Realisation:

D:/MIDI_VERSION

Readme

Walk

Up

Down

Waltz

Float

Slow

Fast

APPENDIX IV - Hand Model Contributors

The author wishes to acknowledge and thank the following individuals for their gracious contribution in the form of providing their left hands for image capture purposes during the data collection phase of the classifier training.

(In no particular order):

Meryl van Noie

Theo Herbst

Larissa Johnson

Marco Titus

Marco Maritz

Samukelo Mhlongo

Sibusiso Tshikili

Nomathamsanqa James

Brandon Andrews

Jason Smythe

Lucy Strauss

REFERENCES

- "computer, n.". 2014. *OED Online*. Oxford: Oxford University Press.
- "esthesis, n.". 2014. *OED Online*. Oxford: Oxford University Press.
- "haptic, n.". 2014. *OED Online*. Oxford: Oxford University Press.
- "Heuristics" 2012. *Encyclopaedia Britannica Online*. Chicago: Encyclopedia Britannica.
- "Ondes Martenot, n.". 2014. *OED Online*. Oxford: Oxford University Press.
- "poiesis, n.". 2014. *OED Online*. Oxford: Oxford University Press.
- "schema, n.". 2014. *OED Online*. Oxford: Oxford University Press.
- "technology, n.". 2014. *OED Online*. Oxford: Oxford University Press.
- Alcorn, P., Mary. 1992. Perspectives of electroacoustic music: a critical study of the electroacoustic music of Jonathan Harvey, Denis Smalley and Trevor Wishart. PhD. Durham University.
- Andreopoulos, A. & Tsotsos, J.K. 2013. 50 Years of object recognition: Directions forward. *Computer Vision and Image Understanding*. 117(8) :827–891.
DOI: <http://dx.doi.org.ezproxy.uct.ac.za/10.1016/j.cviu.2013.04.005>.
- Annamraju, A. 2014. *A Gateway to Robotic Vision – Robotics and Computer Vision*. Available: <http://abhishek4273.com> [2014, February 26].
- Barreiro, D.L. 2010. Sonic Image and Acousmatic Listening. *Organised Sound*. 15(01) :35–42.

Baxter, D., Seagram, R.S. & Amory, A. 2006. The γ Khozi Game Environment: Demonstrations of Two Edventures. *Proceedings of EdMedia: World Conference on Educational Media and Technology 2006*. June. Elaine Pearson & Paul Bohman, Eds. Orlando, FL USA: AACE: Association for the Advancement of Computing in Education. 2513.

Becker, B. & Eckel, G. 1994. *On the Relationship Between Art and Technology in Contemporary Music*. Available:
<http://iem.at/~eckel/publications/becker94c.html> [2012, March 7].

Benjamin, W. 1936. *The Work of Art in the Age of Mechanical Reproduction*. Available:
<http://www.marxists.org/reference/subject/philosophy/works/ge/benjamin.htm> [2014, December 20].

Bhuyan, M.K., Neog, D. & Kar, K. 2012. Fingertip Detection for Hand Pose Recognition. *International Journal on Computer Science and Engineering (IJCSE)*. 4 (3) :501–511.

Blackburn, M. 2011. The Visual Sound–Shapes of Spectromorphology: an illustrative guide to composition. *Organised Sound: An International Journal of Music Technology*. 16 (25 February 2011) :5–13.

Boehmer, K. 2002. Koenig—Sound Composition—*Essay*. In *Electroacoustic Music: Analytical Perspectives*. T. Licata, Ed. USA: Greenwood Press. 59–71.

Bongers, B. 2000. Physical Interfaces in the Electronic Arts: Interaction Theory and Interfacing Techniques for Real-time Performance. In *Trends*

in Gestural Control of Music. M. Wanderley & M. Battier, Eds. Paris: IRCAM. 124–168. Available:

http://www.music.mcgill.ca/~mwanderley/Trends/Trends_in_Gestural_Control_of_Music/; [2012, June 15].

Boulez, P. & Gerzso, A. 1988. Computers in Music. *Scientific American*. 258(4) :44–50.

Brün, H. 1971. *Technology and the Composer*. Available:

<http://www.herbertbrun.org/techcomp.html> [2012, March 13].

Cadoz, C. 1988. Instrumental Gesture and Composition. *Proceedings of the 1988 International Computer Music Conference, Cologne, Germany*. 1988. Ann Arbor, Michigan: Michigan Publishing. 1.

Cadoz, C. & Wanderley, M. 2000. Gesture – Music. In *Trends in Gestural Control of Music*. M. Wanderley & M. Battier, Eds. Paris: IRCAM. 28–65. Available:

http://www.music.mcgill.ca/~mwanderley/Trends/Trends_in_Gestural_Control_of_Music/; [2012, June 15].

Camurri, A., De Poli, G., Friberg, A., Leman, M. & Volpe, G. 2005. The MEGA Project: Analysis and Synthesis of Multisensory Expressive Gesture in Performing Art Applications. *Journal of New Music Research*. 34(1) :5–21. DOI:10.1080/09298210500123895.

Camurri, A. & Moeslund, T., B. 2010. Visual Gesture Recognition. In *Musical Gestures: Sound, Movement, and Meaning*. R. Godøy Inge. & M. Leman, Eds. New York, NY: Routledge. 238–263.

- Camurri, A., Hashimoto, S., Ricchetti, M., Ricci, A., Suzuki, K., Trocca, R. & Volpe, G. 2000. EyesWeb: Toward Gesture and Affect Recognition in Interactive Dance and Music Systems. *Computer Music Journal*. 24(1):57–69.
- Chion, M. 1983. *Guide Des Objets Sonores – Pierre Schaeffer et la recherche musicale*. English translation by John Dack and Christine North, 2009 ed. Paris: Éditions Buchet/Chastel.
- Chung, B. 2013. *Multimedia Programming with Pure Data*. Birmingham, UK. : Packt Publishing Ltd.
- Cogan, R. 1984. *New Images of Musical Sound*. Cambridge, Massachusetts: Harvard University Press.
- Cook, P., R. Ed. 1999. *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. 2nd ed. Massachusetts, USA: MIT Press.
- Cremaschi, A. & Giomi, F. 2004. "Parrrole": Berio's Words on Music Technology. *Computer Music Journal*. 28(1):pp. 26–36.
- Cross, L. 1968. Electronic Music, 1948–1953. *Perspectives of New Music*. 7(1):32–65.
- Dahan, K. & Pollet, D. 2004. Towards an object grammar for electroacoustic music analysis and composition. *Proceedings of the Conference on Interdisciplinary Musicology (CIMO4), Graz, Austria*. 2004. R. Parncutt, A. Kessler & F. Zimmer, Eds. Graz, Austria: University of Graz. np.

- Dahl, S., Bevilacqua, F., Bresin, R., Clayton, M., Leante, L., Poggi, I. & Rasamimanana, N. 2010. Gestures in Performance. In *Musical Gestures: Sound, Movement, and Meaning*. R. Godøy Inge. & M. Leman, Eds. New York, NY: Routledge. 36–68.
- Dahl, S. & Friberg, A. 2007. Visual Perception of Expressiveness in Musicians' Body Movements. *Music Perception: An Interdisciplinary Journal*. 24(5) :433–454.
- Dawson-Howe, K. 2014. *A Practical Introduction to Computer Vision with OpenCV*. Chichester, United Kingdom: John Wiley & Sons Ltd.
- Demers, J. 2010. *Listening Through the Noise: The Aesthetics of Experimental Electronic Music*. New York, New York: Oxford University Press.
- Dimpker, G. 2013. *Extended Notation: The depiction of the unconventional*. Berlin: LIT Verlag.
- Donnarumma, M. 2011. Xth Sense: A Study of Muscle Sounds for an Experimental Paradigm of Musical Performance. *Proceedings of the 2011 International Computer Music Conference, Huddersfield, UK*. 2011. Ann Arbor, Michigan: Michigan Publishing. 243. Available: <http://quod.lib.umich.edu/cgi/p/pod/dod-idx/xth-sense-a-study-of-muscle-sounds-for-an-experimental.pdf?c=icmc;idno=bbp2372.2011.048> [2014, November 20].

- Ellis, C., Adams, T.E. & Bochner, A.P. 2011. Autoethnography: An Overview [40 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*. 12(1):2014.
- Emmerson, S. 2007. *Living Electronic Music*. Surrey, England: Ashgate Publishing Limited.
- Emmerson, S. & Landy, L. 2012. The analysis of electroacoustic music, the differing needs of its genres and categories. *Proceedings of the Electroacoustic Music Studies Network Conference – Meaning and Meaningfulness in Electroacoustic Music, Stockholm, June 2012*. 2012. www.ems-network.org. 1.
- Farnell, A. 2010. *Designing Sound*. Cambridge, Massachusetts, London, England: MIT Press.
- Filimowicz, M. & Stockholm, J. 2010. Towards a Phenomenology of the Acoustic Image. *Organised Sound*. 15(01):5–12.
- Franssen, M., Lokhorst, G. & van de Poel, I. 2013. Philosophy of Technology. In *The Stanford Encyclopedia of Philosophy*. E. Zalta N., Ed. Winter 2013 ed. Stanford University, Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information.
- Garnett, G.E. 2001. The Aesthetics of Interactive Computer Music. *Computer Music Journal*. 25(1):21–33. DOI:10.1162/014892601300126089.
- Gasteiger, B. 2010. A Tangible User Interface for Playing Virtual Acoustics. Diploma. Institute of Electronic Music and Acoustics (IEM) & University of Music and Performing Arts Graz.

- Godøy, R., Inge. 2004. Gestural Imagery in the Service of Musical Imagery. *Gesture-Based Communication in Human-Computer Interaction: 5th International Gesture Workshop, Genova, Italy*. 2003. A. Camurri & G. Volpe, Eds. Berlin Heidelberg: Springer-Verlag. 55.
- Godøy, R., Inge. 2006. Gestural-Sonorous Objects: embodied extensions of Schaeffer's conceptual apparatus. *Organised Sound*. 11 (02):149-157.
- Godøy, R., Inge. 2010. Gestural Affordances of Musical Sound. In *Musical Gestures: Sound, Movement, and Meaning*. R. Godøy Inge. & M. Leman, Eds. New York, NY: Routledge. 104-125.
- Godøy, R., Inge. & Leman, M. Eds. 2010. *Musical Gestures: Sound, Movement, and Meaning*. New York, NY: Routledge.
- Gómez, L. & Degoyon, Y. 2009. Enhancing Pure Data Interactivity with Computer Vision (OpenCV). *3rd PD International Convention*. 2009. Sao Paolo.
- Gritten, A. & King, E. Eds. 2006. *Music and Gesture*. Surrey, England.: Ashgate Publishing Limited.
- Grosshauser, T., GroBekathöfer, U. & Hermann, T. 2010. New Sensors and Pattern Recognition Techniques for String Instruments. *Proceedings of the International Conference on New Interfaces for Musical Expression*. 2010. Sydney, Australia: NIME. 271. Available: http://www.nime.org/proceedings/2010/nime2010_271.pdf [2013, June 29].

- Halmrast, T., Guettler, K., Bader, R. & Godøy, R., Inge. 2010. Gesture and Timbre. In *Musical Gestures: Sound, Movement, and Meaning*. R. Godøy Inge. & M. Leman, Eds. New York, NY: Routledge. 183–211.
- Hamman, M. 2002. From Technical to Technological: The Imperative of Technology in Experimental Music Composition. *Perspectives of New Music*. 40(1):pp. 92–120.
- Hamman, M. 2004. On technology and art: Xenakis at work. *Journal of New Music Research*. 33(2):115–123.
- Hatten, R., S. 2004. *Interpreting Musical Gestures, Topics, and Tropes: Mozart, Beethoven, Schubert*. Bloomington and Indianapolis: Indiana University Press.
- Hatten, R., S. 2006. A Theory of Musical Gesture and its Application to Beethoven and Schubert. In *Music and Gesture*. A. Gritten & E. King, Eds. Surrey, England.: Ashgate Publishing Limited. 1–23.
- Hatten, R., S. 2001. *CSI: Hatten: Musical Gesture*. Available: <http://projects.chass.utoronto.ca/semiotics/cyber/hatout.html> [2013, May 14].
- Hayden, S. & Kanno, M. 2011. Towards Musical Interaction: Sam Hayden's *schismatics* For e-Violin and Computer (2007, REV. 2010). *Proceedings of the 2011 International Computer Music Conference, Huddersfield, UK*. 2011. Ann Arbor, Michigan: Michigan Publishing. 486. Available: <http://quod.lib.umich.edu/cgi/p/pod/dod-idx/towards-musical->

[interaction-schismatics-for-e-violin.pdf?c=icmc;iidno=bbp2372.2011.098](#)

[2014, November 20].

Heidegger, M. 1977. The Question Concerning Technology. In *The Question Concerning Technology and Other Essays*. New York, N.Y. : Harper & Row. 3–35.

Helmholtz, H.v. 1895. *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. 3rd ed. London: Longmans, Green and Co. Available:

<https://archive.org/stream/onsensationsof00helmrich#page/n7/mode/2up>

[2014, December 29].

Helmuth, M. 2006. Barry Truax's *Riverrun*. In *Analytical Methods of Electroacoustic Music*. M. Simoni, Ed. New York, NY: Routledge, Taylor & Francis Group. 187–238.

Hodgkinson, T. 1986. *An Interview with Pierre Schaeffer – Pioneer of Musique Concrète*. Available:

<http://www.timhodgkinson.co.uk/schaeffer.pdf> [2014, February 24].

Impett, J. 1994. A Meta-Trumpet(er). *Proceedings of the 1994 International Computer Music Conference, Denmark*. 1994. Ann Arbor, Michigan: Michigan Publishing. 147. Available:

[http://quod.lib.umich.edu/cgi/p/pod/dod-idx/meta-](http://quod.lib.umich.edu/cgi/p/pod/dod-idx/meta-trumpeter.pdf?c=icmc;iidno=bbp2372.1994.037)

[trumpeter.pdf?c=icmc;iidno=bbp2372.1994.037](#) [2014, November 20].

- Jaimes, A. & Sebe, N. 2007. Multimodal human-computer interaction: A survey. *Computer Vision and Image Understanding*. 108(1-2):116–134.
DOI:<http://dx.doi.org.ezproxy.uct.ac.za/10.1016/j.cviu.2006.10.019>.
- Jenseniuss, A., Wanderley, M., Godøy, R. & Leman, M. 2010. Musical Gestures: Concepts and methods in research. In *Musical Gestures: Sound, Movement, and Meaning*. R. Godøy Inge. & M. Leman, Eds. New York, NY: Routledge. 12–35.
- Jessop, E. 2010. A Gestural Media Framework: Tools for Expressive Gesture Recognition and Mapping in Rehearsal and Performance. Master Of Science. MIT.
- Karam, M. 2006. A Framework for Research and Design of Gesture-Based Human Computer Interactions. PhD. University of Southampton.
- Kimura, M. & Rasamimanana, N. e. a. 2012. Extracting Human Expression For Interactive Composition with the Augmented Violin. *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*. 2012. Ann Arbor, Michigan: University Of Michigan.
- Kimura, M. & Risset, J. 2006. Auditory Illusion and Violin: Demonstration of a Work by Jean-Claude Risset Written for Mari Kimura. *Proceedings of the 2006 International Conference on New Interfaces for Musical Expression (NIME06)*. 2006. Paris, France: 407. Available: http://www.nime.org/proceedings/2006/nime2006_407.pdf [2014, May 20].
- Kojs, J. 2005. Piano Case, Keyboard, And Strings: Separation-Derived Musical Instruments In An Interactive Composition. *Proceedings of the*

- 2005 International Computer Music Conference, Barcelona, Spain. 2005.*
Ann Arbor, Michigan: Michigan Publishing. n.p. Available:
<http://quod.lib.umich.edu/cgi/p/pod/dod-idx/piano-case-keyboard-and-strings-separation-derived-musical.pdf?c=icmc;idno=bbp2372>. 2005. 194
[2014, November 20].
- Kreidler, J. 2013. *Loadbang: Programming Electronic Music in Pure Data*.
2nd ed. Hofheim, Germany: Wolke Verlag.
- Kühl, O. 2011. The Semiotic Gesture. In *New Perspectives on Music and Gesture – SEMPRE studies in the psychology of music*. A. Gritten & E. King, Eds. Surrey, England: Ashgate Publishing Limited. 124–129.
- Laden, B. & Keefe, D.H. 1989. The Representation of Pitch in a Neural Net Model of Chord Classification. *Computer Music Journal*. 13(4):12–26.
DOI:10.2307/3679550.
- Landy, L. 2006. The Intention/Reception Project. In *Analytical Methods of Electroacoustic Music*. M. Simoni, Ed. New York, NY: Routledge, Taylor & Francis Group. 29–53.
- Laske, O. 1975. *Introduction to a generative theory of music*. Utrecht State University, Institute of Sonology.
- Laske, O. 1988. Introduction to Cognitive Musicology. *Computer Music Journal*. 12(1):43–57. DOI:10.2307/3679836.
- Laske, O. 1999. *Furies and Voices: Composition-Theoretical Observations*. In *Otto Laske: Navigating New Musical Horizons*. J. Tabor, Ed. Westport: Greenwood Press. 151–167.

- Leman, M. & Godøy, R., Inge. 2010. Why Study Musical Gestures? In *Musical Gestures: Sound, Movement, and Meaning*. R. Godøy Inge. & M. Leman, Eds. New York, NY: Routledge. 3–11.
- Leman, M. 2008. *Embodied music cognition and mediation technology*. Cambridge: MIT Press Cambridge, MA.
- Leman, M. 1999. Adequacy Criteria for Models of Musical Cognition. In *Otto Laske: Navigating New Musical Horizons*. J. Tabor, Ed. Westport: Greenwood Press. 93–120.
- Licata, T. Ed. 2002. *Electroacoustic Music: Analytical Perspectives*. USA: Greenwood Press.
- Livingston, P. 2013. History of the Ontology of Art. In *The Stanford Encyclopedia of Philosophy*. E. Zalta N., Ed. Summer 2013 ed. Stanford University, Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information.
- Loy, G. 2006. *Musimathics: the mathematical foundations of music. Volume 1*. 2011th ed. Cambridge, MA.: MIT Press.
- Machover, T. 1992. *Hyperinstruments: A Progress Report, 1987–1991*, MIT Media Lab. Massachusetts: MIT.
- MacRitchie, J., Buck, B. & Bailey, N.J. 2013. Inferring musical structure through bodily gestures. *Musicae Scientiae*. 17(1):86–108.
- Maestre, E. 2009. Modelling instrumental gestures: An analysis/synthesis framework for violin bowing . PhD. Universitat Pompeu Fabra, Spain.

- Martin, J.L. 2014. Semiotic resources of music notation: Towards a multimodal analysis of musical notation in student texts. *Semiotica*. 2014 (200) :185–201. DOI:10.1515/sem-2014-0006.
- May, A. 2006. Philippe Manoury's *Jupiter*. In *Analytical Methods of Electroacoustic Music*. M. Simoni, Ed. New York, NY: Routledge, Taylor & Francis Group. 145–185.
- Mazzoli, M. 2014. Emerging Musical Structures: A Method for the Transcription and Analysis of Electroacoustic Music. PhD. City University of New York.
- McNutt, E. 2003. Performing Electroacoustic Music: A Wider View of Interactivity. *Organised Sound*. 8 (3) :297–304.
- Meric, R. 2011. "Music is not a language..." Listening to Xenakis's Electroacoustic Music. *Proceedings of the Xenakis International Symposium, Southbank Centre, London*. 2011. 1. Available: <http://www.gold.ac.uk/media/05.2%20Renaud%20Meric.pdf> [2014, May, 19].
- Metzer, D. 2009. *Musical Modernism at the Turn of the Twenty-First Century*. Reprinted 2011 ed. New York, NY: Cambridge University Press.
- Nattiez, J.J. 1990. *Music and discourse : toward a semiology of music*. Princeton, N.J. : Princeton University Press.
- Newell, A. & Simon, H., A. 1972. *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.

- OpenCV Development Team 2014. *The OpenCV Reference Manual – Release 2.4.9.0*. Available: <http://docs.opencv.org/opencv2refman.pdf> [2014, June 2].
- Oritz, M. 2012. A Brief History of Biosignal-Driven Art: From Biofeedback to Biophysical Performance. *Electronic Journal of Electroacoustics*. 14 (2):2014, November 20.
- Overholt, D., Thompson, J., Putnam, L., Bell, B., Kleban, J., Sturm, B. & Kuchera-Morin, J. 2009. A Multimodal System for Gesture Recognition in Interactive Music Performance. *Computer Music Journal*. 33 (4):69–82.
- Papageorgiou, C.P., Oren, M. & Poggio, T. 1998. A general framework for object detection. *Computer Vision, 1998. Sixth International Conference On*. 555.10.1109/ICCV.1998.710772.
- Pierce, J. 1999. Hearing in Time and Space. In *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. P. Cook R., Ed. 2nd ed. Massachusetts, USA: MIT Press. 89–103.
- Pierce, J. 1999. Introduction to Pitch Perception. In *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. P. Cook R., Ed. 2nd ed. Massachusetts, USA: MIT Press. 57–70.
- Puckette, M. 1988. The Patcher. *International Computer Music Conference Proceedings*. 1988. Ann Arbor, Michigan: Michigan Publishing. 420.
- Puckette, M. 1997. Pure Data. *Proceedings of the 1997 International Computer Music Conference Proceedings, San Francisco*. 1997. Ann Arbor, Michigan: Michigan Publishing. 269. Available:

<http://quod.lib.umich.edu/cgi/p/pod/dod-idx/pure-data.pdf?c=icmc;idno=bbp2372.1997.060> [2012, March 20].

- Puckette, M. 2007. *The Theory and Technique of Electronic Music*.
Singapore: World Scientific Publishing Company.
- Ramberg, B. & Gjesdal, K. 2014. Hermeneutics. In *The Stanford Encyclopedia of Philosophy*. E. Zalta N., Ed. Winter 2014 ed. Stanford University, Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information.
- Rasamimanana, N. 2004. *Gesture Analysis of Bow Strokes Using an Augmented Violin*. Available: <http://recherche.ircam.fr/equipes/temps-reel/movement/publications/NicolasRasamimanana-master.pdf> [2014, June 18].
- Read, G. 1979. *Music Notation: A Manual of Modern Practice*. 2nd ed. New York, NY: Taplinger Publishing Co., Inc.
- Reimer, B. 1964. Information Theory and the Analysis of Musical Meaning. *Bulletin of the Council for Research in Music Education*. (2):14–22. DOI:10.2307/40375245.
- Rink, J., Spiro, N. & Gold, N. 2011. Motive, Gesture and the Analysis of Performance. In *New Perspectives on Music and Gesture – SEMPRE studies in the psychology of music*. A. Gritten & E. King, Eds. Surrey, England: Ashgate Publishing Limited. 267–292.
- Risset, J. 2002. Foreword. In *Electroacoustic Music: Analytical Perspectives*. T. Licata, Ed. USA: Greenwood Press. xii–xviii.

- Rizzolatti, G. & Luppino, G. 2001. The Cortical Motor System. *Neuron*. 31 (6) :889–901. DOI:[http://dx.doi.org.ezproxy.uct.ac.za/10.1016/S0896-6273\(01\)00423-8](http://dx.doi.org.ezproxy.uct.ac.za/10.1016/S0896-6273(01)00423-8).
- Rosenboom, D. 1997. *Extended Musical Interface with the Human Nervous System*. Available: http://davidrosenboom.com/sites/default/files/media/downloads/MusInter_LEO_97.final_w_figs.pdf [2014, November 23].
- Rowe, R. 1993. *Interactive Music Systems*. United States of America: MIT Press.
- Schneider, A. 2010. Music and Gestures: A Historical Introduction and Survey of Earlier Research. In *Musical Gestures: Sound, Movement, and Meaning*. R. Godøy Inge. & M. Leman, Eds. New York, NY: Routledge. 69–100.
- Seo, N. 2007. *Tutorial: OpenCV haartraining (Rapid Object Detection with a Cascade of Boosted Classifiers Based on Haar-like Features)*. Available: <http://note.sonots.com/SciSoftware/haartraining.html> [2012, January 10].
- Settel, J. 2001. The use of real-time interactive music systems in music composition and performance. University Microfilms International (UMI) Ann Arbor, MI.
- Shepard, R. 1999. Cognitive Psychology and Music. In *Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics*. P. Cook R., Ed. 2nd ed. Massachusetts, USA: MIT Press. 21–35.

- Siewert, C. 2011. Consciousness and Intentionality. In *The Stanford Encyclopedia of Philosophy*. E. Zalta N., Ed. Fall 2011 ed. Stanford University, Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information.
- Simoni, M. Ed. 2006. *Analytical Methods of Electroacoustic Music*. New York, NY: Routledge, Taylor & Francis Group.
- Smalley, D. 1997. Spectromorphology: explaining sound-shapes. *Organised Sound*. 2 (02) :107–126.
- Smith, E., E. & Kosslyn, S., M. 2006. *Cognitive Psychology: Mind and Brain*. 1st ed. Upper Saddle River, N. J. : Pearson/Prentice Hall.
- Snow, C., P. 1959. *The Two Cultures and The Scientific Revolution*. Seventh Printing, 1961 ed. New York, N. Y. : Cambridge University Press.
- Sonami, L. 2010. *Lady's Glove: Laetitia Sonami*. Available: <http://sonami.net/works/ladys-glove/> [2014, November 23].
- Strange, P. 2001. *The Contemporary Violin: Extended Performance Techniques*. Berkeley, California: University Of California Press.
- Tabor, J. 1999. A Pioneer in Composition and Research. In *Otto Laske: Navigating New Musical Horizons*. J. Tabor, Ed. Westport: Greenwood Press. 1–18.
- Tarabella, L. & Bertini, G. 2002. Wireless technology in gesture controlled computer generated music. *Workshop on Current Research Directions in Computer Music, Barcelona, Spain*. 2001. C.L. Buyoli & R. Loureiro, Eds. Barcelona, Spain: Pompeu Fabra University. 102.

Available:

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.5.2424&rep=rep1&type=pdf> [2014, November 22].

- Teruggi, D. 2007. Technology and musique concrète: the technical developments of the Groupe de Recherches Musicales and their implication in musical composition. *Organised Sound*. 12(03):213–231.
- Thagard, P. 2014. Cognitive Science. In *The Stanford Encyclopedia of Philosophy*. E. Zalta N., Ed. Fall 2014 ed. Stanford University, Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information.
- Thomas, N., J.T. 2014. Mental Imagery. In *The Stanford Encyclopedia of Philosophy*. E. Zalta N., Ed. Fall 2014 ed. Stanford University, Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information.
- Tonder, v.C. 2004. Music Composition and Performance in Interactive Computer/Human Systems. Honors. University of Witwatersrand.
- Truax, B. 1999. Sonology: A Questionable Science Revisited. In *Otto Laske: Navigating New Musical Horizons*. J. Tabor, Ed. Westport: Greenwood Press. 21–36.
- Vaggione, H. 2001. Some Ontological Remarks about Music Composition Processes. *Computer Music Journal*. 25(1):54–61.
DOI:10.1162/014892601300126115.

- Viola, P. & Jones, M. 2001. Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 511. 10. 1109/CVPR. 2001. 990517.
- Waisvisz, M. 1985. The Hands: A Set of Remote MIDI-Controllers. *Proceedings of the 1985 International Computer Music Conference, Simon Frasier University, Canada*. 1985. Ann Arbor, Michigan: Michigan University. 313. Available: <http://quod.lib.umich.edu/cgi/p/pod/dod-idx/hands-a-set-of-remote-midi-controllers.pdf?c=icmc;idno=bbp2372.1985.049> [2014, November 20].
- Wanderley, M. & Battier, M. Eds. 2000. *Trends in Gestural Control of Music*. Paris: IRCAM. Available: http://www.music.mcgill.ca/~mwanderley/Trends/Trends_in_Gestural_Control_of_Music/ [2012, June 15].
- Wanderley, M.M. & Orio, N. 2002. Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI. *Computer Music Journal*. 26(3, New Performance Interfaces):62-76.
- Warrington, M. 2005a. A preliminary investigation into interactive computer game music. M.Mus. University of Kwa Zulu-Natal, Durban.
- Warrington, M. 2005b. The burning ground: a portfolio of compositions with music and sound design for "gammaKhozi" the VLS game. M. Mus. University of Kwa Zulu-Natal, Durban.

- Warrington, M. & Annamraju, A. 2014. *OpenCV Questions and Answers Forum*
[The number of negative images must be greater than twice the number
of positive images.]. Available:
[http://answers.opencv.org/question/29126/haarcascade-finger-tip-
detection-and-image-prep/?comment=29127#comment-29127](http://answers.opencv.org/question/29126/haarcascade-finger-tip-detection-and-image-prep/?comment=29127#comment-29127) [2014, February
26].
- Welch, E.P. 1939. *Edmund Husserl's Phenomenology*. California: University of
California Press.
- Wertheimer, M. 2014. Music, thinking, perceived motion: The emergence of
Gestalt theory. *History of Psychology*. 17(2):131-133.
DOI:10.1037/a0035765.
- Winkler, T. 1995. Making Motion Musical: Gesture Mapping Strategies for
Interactive Computer Music. *Proceedings of the 1995 International
Computer Music Conference*. 1995. Ann Arbor, Michigan: Michigan
Publishing. 261.
- Young, D. 2007. A methodology for investigation of bowed string
performance through measurement of violin bowing technique. PhD. MIT,
Cambridge, MA.
- Zhang, Y. & Candy, L. 2007. *An in-depth case study of art-technology
collaboration*. Washington, DC, USA: ACM.