Implicit Relevance Feedback in Interactive Music: Issues, Challenges, and Case Studies

Lars Graugaard
lag@cs.aaue.dk
Department of Software and Media Technology
Aalborg University Esbjerg – Denmark

ABSTRACT

This paper presents methods for correlating a human performer and a synthetic accompaniment based on Implicit Relevance Feedback (IRF) using Graugaard's expanded model for interactive music (Graugaard 2006c). The research is the result of experience with practical work with interactive music systems developed 2004-06 for a body of commissioned works and is based on human perception of music as an expressive artform where musically significant data may be present not only in the audio signal but also in human gestures and in physiological data. The relevance and feasibility of including expression and emotion as a high-level signal processing means for bridging man and machine is discussed. The resulting model is multi-level (physical, sensorial, perceptual, formal, expressive) and multi-modal (sound, human gesture, physiological), which makes it applicable to purely musical contexts, as well as intermodal contexts where music is combined with visual and/or physiological data.

KEY AREAS

Task-based IIR and information behavior in media and genre-dependent applications; personalized and collaborative information access in context.

1 INTRODUCTION

Music performance is an artform where information is channeled bi-directionally through primarily the auditory and secondarily the haptic sense. The abstraction level is high, and the nature of music as a performed art requires utmost precision and speed in information processing for immediate and intuitive usage in the actuate-sense loop (Fig. 1-1). Interactive music is the performance with computers where a digital accompaniment is generated in real-time based on the combination of a digital score, a performance score, and the digital systems analysis of the human performance for immediate auditory feedback into the actuate-response loop (Fig. 1-2). In both cases the nature and relevance of the obtained information varies throughout a composition according to the requirements of its 'good performance'.

The term *interactive* describes the functionality of the computer rather than the music itself. Music is in its nature interactive being a performing art, and the task of the composer includes programming for such interaction while observing that it takes place in accordance with the compositional argument of the work. All relevant musical information imparted by the composer is no longer exclusively contained in the performance instructions because the interacting computer takes independent part in the music performance, capable of affecting its low, mid, and high levels (Graugaard 2004a). The interdependency of these two instruction sets defines the nature and content of the musical work, its limitations in execution, and its characteristics in performance.

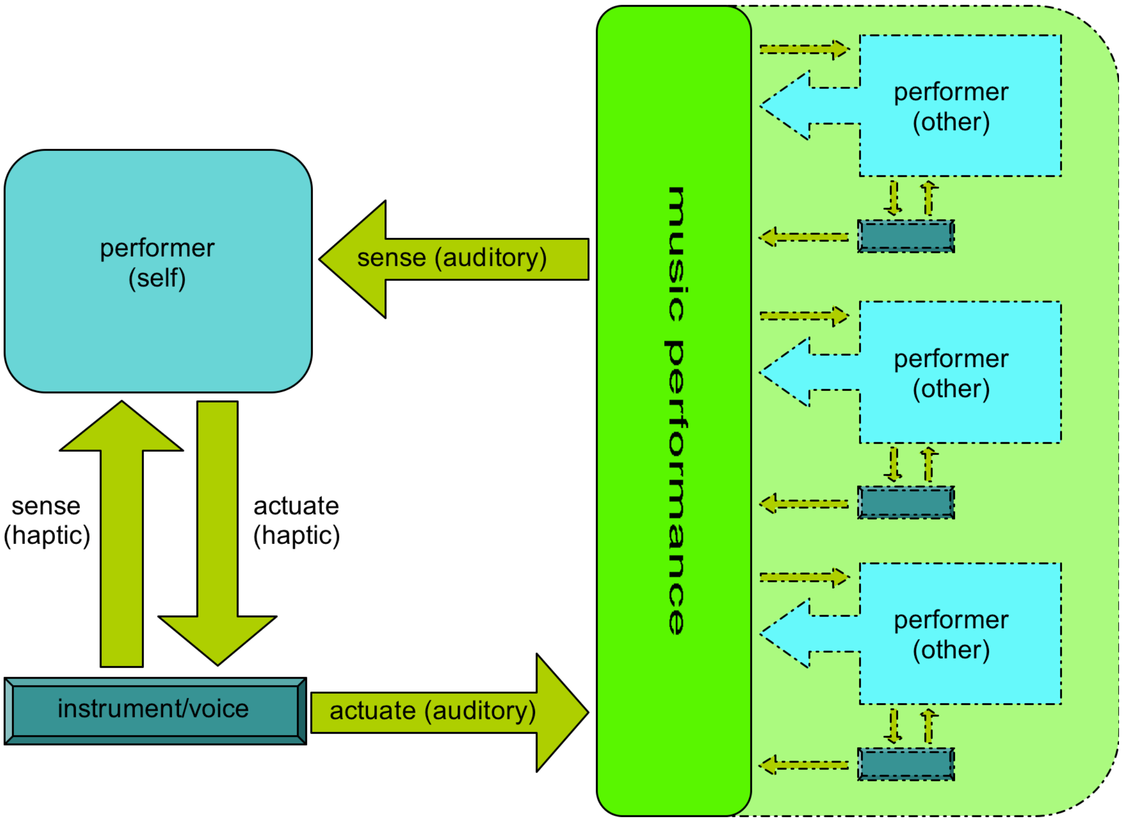

Fig. 1-1 *Auditory and haptic actuate-sense flow during the act of music performance from perspective of self.*

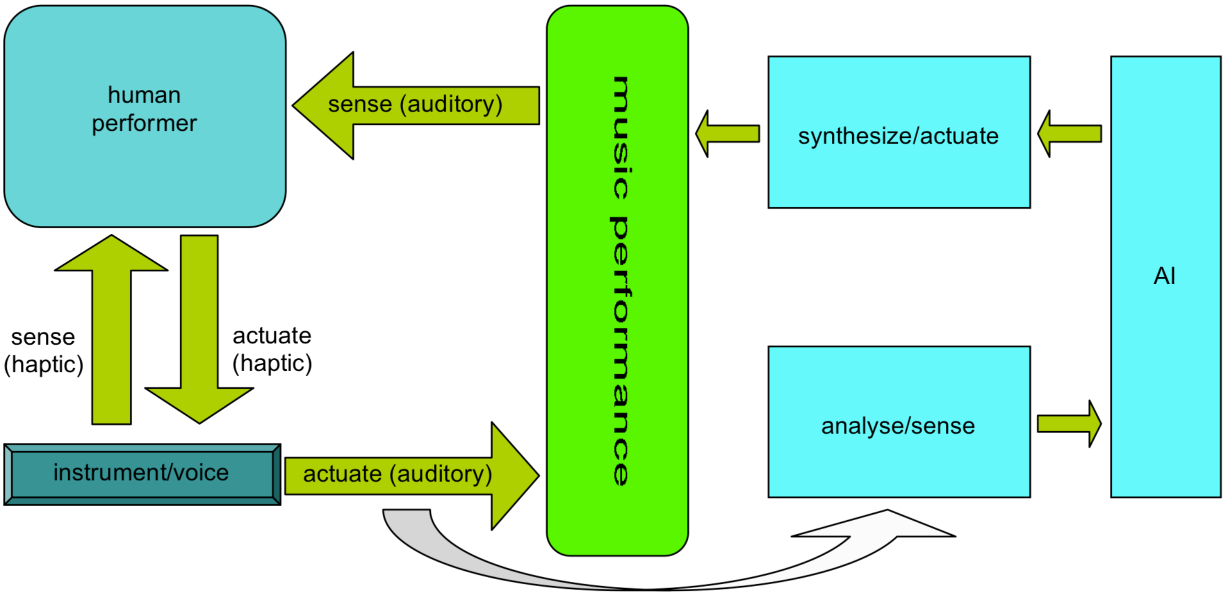

Fig. 1-2 *Actuate-response loop in interactive music.*

2 *LA QUINTRALA* – AN INTERACTIVE OPERA

Opera is a musical drama where feelings of love, longing, hope, fear, and hatred are the essence of tragedy and drama. To perform their parts the opera singers require an accompaniment which is both expressive (enhancing their dramatic presence) and functional (providing tonal orientation). The electronic accompaniment of an interactive opera such as *La Quintrala* (Graugaard 2004b) must consequently be able to relate the pitches of the singers to the chord succession underlying the music, render the singers' expression, and give the composer adequate freedom for compositional development.

2.1 *NOTATED AND ELECTRONIC SCORE*

The notated score of *La Quintrala* contains the parts for the singers and the accompanying chords as a sequence of basic pitch-class sets. The electronic score is an application containing all methods needed for voice analysis and comparison, methods for sound synthesis and diffusion, and all procedural methods for interaction. An important feature of the system is that the voice-to-chord comparison is independent of the sound synthesis. The advantage is that a uniform and consistent way of communicating basic harmonic support to the singers can be maintained, while the characteristics of the accompaniment may vary. Notation of the vocal parts is done in traditional fashion for the singers' best reading (Fig. 2-1).

2.2 *INTERACTIVE ACCOMPANIMENT*

La Quintrala does not require a conductor. Instead, it relies on the auditory channel for passing all music performance information: the singers perform the notated musical phrases while listening for the context, simultaneously process information concerning intonation and expression, and continuously adjusts their interpretation. The reference to the aquired information is the dramatic and musical properties at that particular point and their expected development, and this may involve the efforts of the other singers. The information to be managed is highly contextual, particulary when considering that music in western musical notation system consists of just 12 different pitches (the vocal score of *La Quintrala* covers app. 4 octaves) and a dynamic range of 8 loudness categories.

2.2.1 AUDIO ANALYSIS AND HARMONIC SUPPORT

Harmonic support is based on reducing the voice analysis to three primary frequency components, compare them to the prevailing chord and forward the closest match for the sound synthesis step. Closest match is dynamically defined by the compositional context and is regulated by the harmonicity factor (the ratio between the input and the candidates) and affected by the dissonance factor (for microtonal deviation) in a continuum from no correlation to complete chordal alignment. The musical pitches are excited by the singers in accordance with their performance (the voice analysis and the parameter settings) and are restricted to those contained in the present chord. The pitches of the chord are not altered by the singers, and the singers are restricted in performance by the notation of the vocal score.

As a consequence, the accompaniment is always the same in terms of overall harmonic content and purpose, but the particular rendering varies according to the performance. The nature of the music is therefore quite similar to traditional music which is also characterized by slight differences between performances, yet always within the limits defined by the score. The main functional difference is that interactive music provide for performer influence on the electronic accompaniment through the performer's manipulation of the musical aspects of timbre, time, and dynamics. This is a fundamentally different situation for the performer in respect to acoustic music performance, and brings with it unique responsibilities and possibilities.

__Fig. 2-1__ La Quintrala – score notation example.

2.2.2 EXPRESSION ANALYSIS AND PRESENCE ENHANCEMENT

Musical expression takes place through minute manipulations of timbre, dynamics, intonation, and timing. Musical interpretation involves continuous modification of these elements and requires substantial insight and experience with performance practises in different musical styles and epochs. Presence enhancement in *La Quintrala* – the enlargement of human expression by a digital component – relies on mapping the singers' such musical and dramatic efforts onto the larger canvas of the electronic accompaniment. The singers may not feel that they actually control the accompaniment (even though it is highly influenced by them) but they always intuitively feel that their musical efforts are supported by the sound synthesis and that their expressive intentions are augmented, juxtaposed, and commented on by the accompaniment, and hereby underscoring the development and dramatic moments of their role.

2.2.3 AUDITORY INFORMATION FEEDBACK

Interaction in music takes place in the time continuum from local triggering to global implicit affect (Lippe 1996) in a mapping continuum from one-to-one to one-to-many/many-to-one (Wanderley 2000) and in a structural continuum from signal level to gesture level. Using Lesaffre's feature taxonomy for content based audio queries (Lesaffre et al. 2003) the feedback is seen to reside in the physical, sensorial, perceptual, formal, and expressive concept level depending on context and time-frame of the analysis (Fig. 2-2). This means that all levels of the musical work is available for information feedback even though the lower sensorial and percepual levels are those most relevant to the singers.

STRUCT		CONCEPT LEVEL		MUSICAL CONTENT FEATURES				
CONTEXTUAL	global beyond 3 sec	HIGH II	EXPRESSIVE	cognition | emotion | affect = *syntactic+semantic concepts*				
				melody	harmony	rhythm	source	dynamics
		HIGH I	FORMAL	key profile	tonality cadence	rhythmic patterns tempo	instrument voice	trajectory articulation
	global < 3 sec	MID	PERCEPTUAL	successive intervallic pattern	simultane intervallic pattern	beat IOI	spectral envelope	dynamic range sound level
				pitch		time	timbre	loudness
NON-CONTEXTUAL	local + spatial	LOW II	SENSORIAL	periodicity pitch pitch deviations fundamental frequency		note duration onset offset	roughness spectral flux spectral centroid	neural energy peak
	local + temporal	LOW I	PHYSICAL					
				frequency		duration	spectrum	intensity

Fig. 2-2 *Lessafre's taxonomy for musical content features.*

2.2.4 IMPLICIT RELEVANCE OF THE AUDITORY FEEDBACK

The musical information provided by the interactive accompaniment concerning correctness of intonation and enhancement of expression corresponds to two different categories of feedback.

2.2.4.1 TONAL CONTEXT FEEDBACK

The tonal context of the feedback supports the continuous intonation correction that a singer does when performing. The relevance of the feedback is governed by the present chords. The feedback is a tonal guide, and the singers use it to align and expressively vary their intonation and timing at moments of chordal transitions. It is in essence a constant feedback to the singer upon which the singer has no influence. The system aims to inform the singer of the intonational correctness, prevailing tonal orientation, and suggests future intervalic intonation. The usage of the information has no influence on the system.

2.2.4.2 EXPRESSION RELEVANCE FEEDBACK

Modulating musical expression is essential to the singers in articulating the development over time of their performance. Being shapeable – and therefore requiring shaping – it is a panorama where composer and singers share requirements for the good interpretation of the music, albeit from different perspectives: the composer needs the singers to render expressivity to the music, while the singers needs the composer to prepare an optimal platform for their stage presence. It is furthermore essential for the composer to shape the opera on the highest formal level in articulating the dramatic development of the opera's fundamental argument, which makes his dependency on the efforts and abilities of the singers even more acute.

2.2.4.3 THE MEANING OF THE FEEDBACK

Concepts from interactive, narrative storytelling from multimedia are tempting to apply to the interactive accompaniment in *La Quintrala*, because they also deal with organised time structures carrying some form of meaning explored in real-time by a user (the singer). But there are fundamental differences between music and narrative and storytelling – primarily because music is not directly comparable to language, since its units do not carry absolute meaning – making a cautious approach necessary. The lack of absolute meaning makes the intuitive link between singer and accompaniment very important, because the passing of IRF will depend on their degree of symbiotic effort-feedback relationship (Fig. 2-3). The consequence for the singers is that their musical intensities and the intentions inherent in their singing is projected onto the accompanying sound canvas, hereby supporting their musicality and enhancing their stage presence. The consequence for the composer is that the real-time musical and dramatic efforts can be included as an artistic dimension directly affecting musical output.

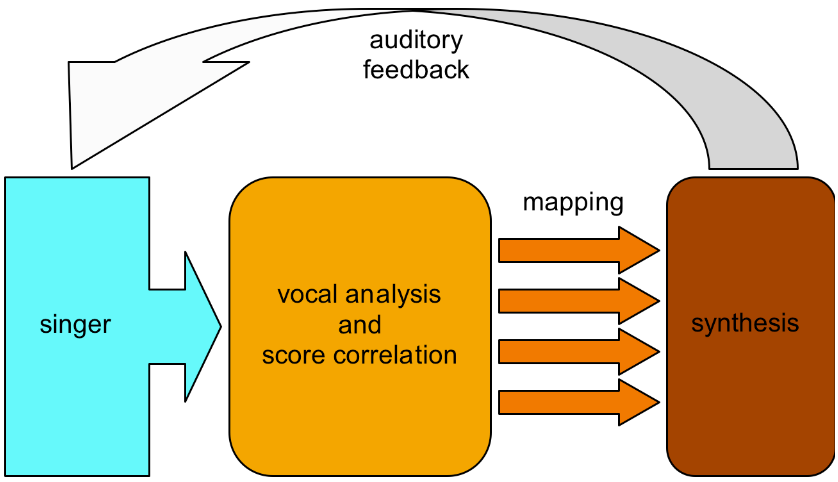

Fig. 2-3 *Effort-feedback chain in* La Quintrala.

2.2.5 SHAPEABLE RELEVANCE

The implicit relevance of the feedback is entirely decided by the composer in accordance with the compositional needs. The decisions include methods of sound synthesis, pitch content (simultaneous and sequential), time segmentation as rhythms, etc. It can also be described as the digital score's 'composition space' where the composer articulates the digital score that will render the electronic accompaniment. Conceptually being the digital counterpart to the paper score, it is in practical terms very different since it is articulated through programming all methods used in sound analysis, mapping, parameterization, and sound synthesis including sound diffusion and temporal cross-relations. The artistic result of this 'shapeable relevance' is the composer's ability to guide the singers through the tonal and emotional soundscape of the opera while projecting their musical efforts onto it, bringing the opera's drama to life and communicating itself to the audience. The parameters for the accompaniment are the composer's maneuvering inlets into the electronic soundscape, once a scene's fundamental methods have been decided.

La Quintrala's triggering system for the accompaniment takes input from four different places, individually or in combination, and each with appropriate parameters (Fig. 2-4). The triggering filters control density of the data passed the comparer and control the activity level and responsivenes of the accompaniment. The mapping incorporates and guides the expressive qualities of the singing. The parameters exposed provide a global and unified means of controlling affection possibilities whereby the system responds to compositional choices stored at system level, to expressive parameters from the singers, and to real-time interpretative relations developed on an *ad hoc* basis between singers and sound synthesis.

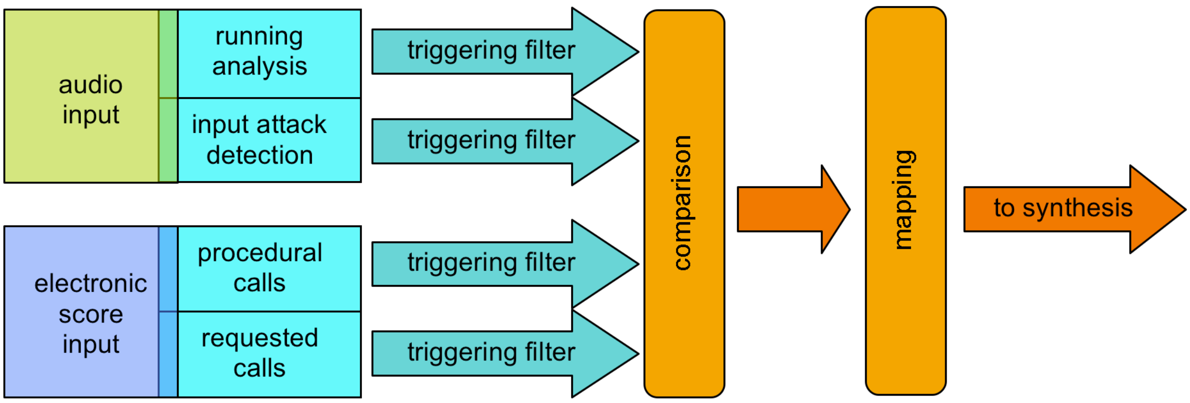

Fig. 2-4 *Flow structure of* La Quintrala*'s shapeable triggering system.*

As an example, a particular case of IRF concerns issues of timing and event coincidence. *La Quintrala* generates precise, coordinated events in real-time by filtering the performance analysis affect to the chosen synthesis method(s). This means that the performer defines the perceived rhythm periodicity when the rhythm texture is sufficiently complex to be ambiguous in itself, yet sufficiently simple to maintain the feeling of pulse. As a consequence the performer 'defines' the overall periodicity through the periodicity of the sung melismas.

La Quintrala was premiered September 2nd 2004 at Den Anden Opera, Copenhagen, Denmark, with a subsequent 22 performances in Denmark, Germany, and Sweden.

3 AN ENHANCED CONTEXT OF GESTURE AND EMOTION

Computer music is highly interdisciplinary and closely related to developments in computer hard- and software, digital signal analysis, artificial intelligence, psychoacoustics, perception, and music cognition (Moore 1990). Areas that recently have been drawn into the disciplinary context of interactive music are interactive storytelling and emotion analysis and synthesis, and if Moore's diagram for the disciplinary context of computer music is modified accordingly and targeted at interactive music, a complex web of highly sophisticated areas going into the making of interactive music appears (Graugaard 2006c, fig. 3-1). As these areas develop, so does the composer-programmers' possibility to apply and explore unique and distinct musical propositions.

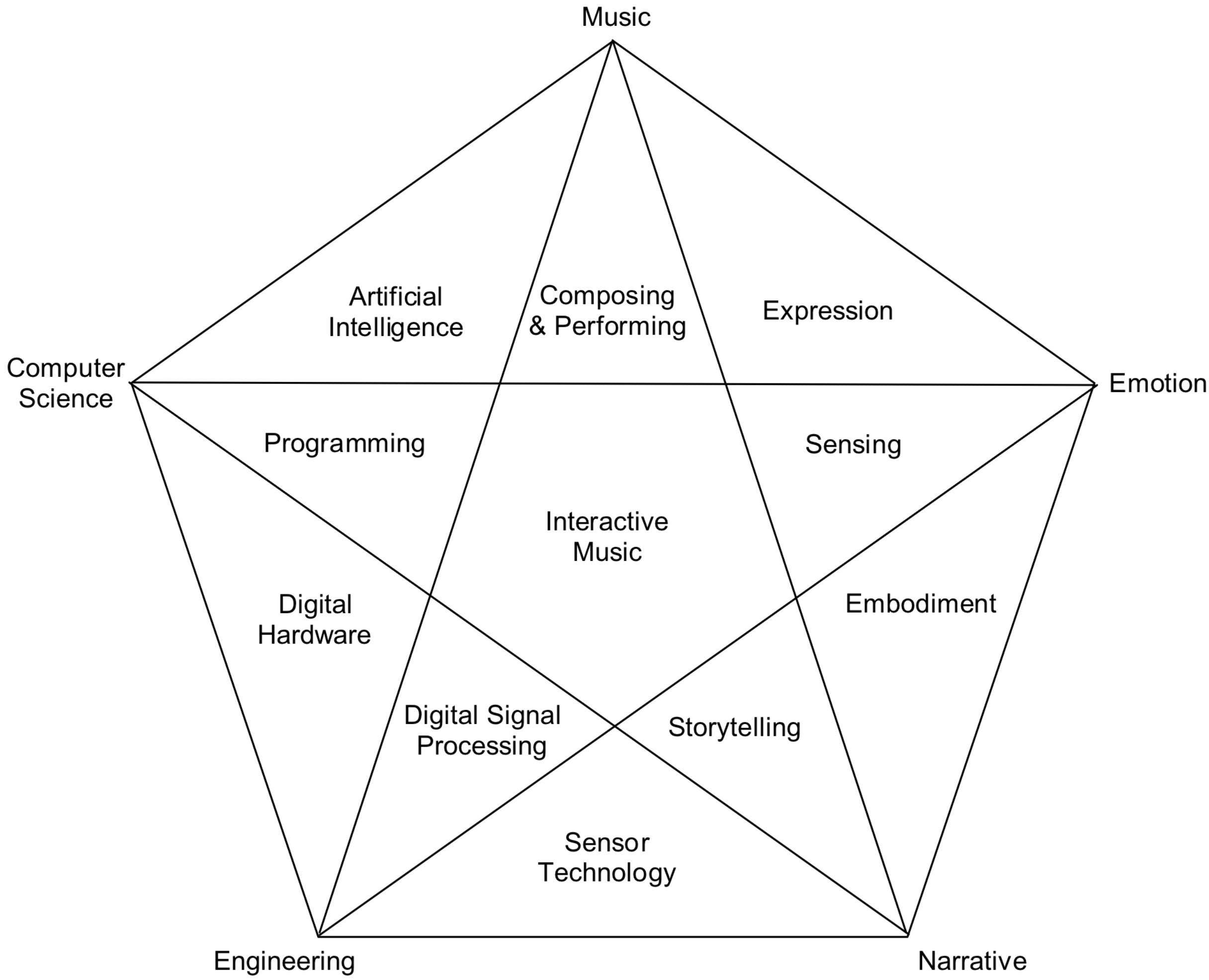

Fig. 3-1 *Disciplinary context of interactive music.*

3.1 EXPRESSION, GESTURE, AND EMOTION IN MUSIC

Notation and appreciation of music do not seem to have much in common. Traditional music listening models based on notes grouped into melodies, rhythms, chords, and harmonic progressions are only applicable to a small group of expert listeners such as musicians, and there is strong evidence that non-musicians do not hear music in these terms (Martin et al. 1998). But emotional skills are an essential part of human intelligence; they modulate human communication and are fundamental to human activities including the making and consumption of music (Scherer et al. 2001). The significance of the emotional aspects in musical content is readily appreciable by scholar and layman alike and it seems obvious to capacitate computers with the ability to recognise and respond to human emotion in the field of interactive music performance.

The strength of emotion and the ease with which humans detect basic emotions in music have contributed to establishing sets of detectable expression categories through a limited number of high-level music parameters (Sundberg 1993, Gabrielsson et al. 1996, Bresin et al. 2000, Juslin 2000, Canazza et al. 2001, Zanon et al. 2003, Canazza et al. 2004). In (Camurri et al. 2000), the authors report on an abstract space representing music's basic emotion content. The emotions are derived through analysis of perceptual content of a musical statement played with different expressive intentions and the detection of emotions are sufficiently solid to be independent of the instrument on which the music was performed. The emotion space is then used for synthesising expression in a mechanical performance of the same musical statement, where emotion descriptors are mapped to high-level music expression parameters for microvariations in timing, dynamics, and pitch.

4 IRF IN GESTURE AND EMOTION DETECTION

The origins of music cannot be easily separated from those of dance – stomping, clapping hands, snapping fingers and enjoying the movement of one's body is part of the earliest form of revered feelings – as archetypal forms were laid down by early humanoids through stylised movements accompanied by 'body music'. This complementary nature of dance and music is as vivid as ever, as music performance displays strong ties with complimentary physical gestures for conveying musical expression. Such is the strength of our conditioned link between gesture and sound that Denis Smalley bases his concept of gestural surrogates on the perceived link between a musical gesture and an underlying human or instrumental gesture (Smalley 1997). Analysis of a performer's gestures for expressive cues and emotion content are reported in (Zannos et al. 1997, Camurri et al. 1999 and 2002, Friberg et al. 2002, Friberg 2004).

Correlations between the perceived emotion content of a musician's performance movements and the analysed emotion content of the music performed is studied in (Dahl et al. 2004), and low-level physiological signals is seen as a non-linguistic description that unites perception with its sensimotoric basis as emotion identifiers (Leman 2006). Examples of such data are a performer's physiological signals transmitting emotion state information. This opens up to a multimodal field as relevant for the interaction loop of audio, gesture, and physiological signals. Such context dependent interfaces for capturing human expression uses an amalgam of the present state and intended trajectory of the performance, where simple bodily expression is the source of human expressivity when emotion-related physical energy translates into different forms of physical gestures.

4.1 HUMAN EXPRESSION CAPTURING

Gestures drawn with a light-pen on a wall detected through standard computer vision techniques are used as a robust musical interface for real-time adaptable control of many musical gestures in *D^2MMG* (Jensen et al. 2006). The drawing gestures are structured into primary and secondary gestures and affect low- and high-level aspects of sound and music. This enables expressive live performance and improvisation with sequence creation where musical elements such as timbre, envelope, duration, vibrato, pitch, tonality, chromaticity, rhythm patterns, and melodical sequences can be affected in real-time. High-level aspects such as note density and chordal and time alignment are procedurally controlled.

The audiovisual composition *Gestures You Made* (Graugaard 2006a) for oboe and interactive computer is largely interactive through sound analysis but towards the end of the piece the player stops playing at moments, and instead affects the sound synthesis by assuming body positions and making physical gestures inside a 'videospace' adjacent to the performance space (Graugaard 2006b). The result can be likened to an externalized monologue, where the performer comments on own playing through physical gestures, for then to comment on these gestures through playing. The link between the two affect channels is the parallelism between expression in sound and in human gesture. The parallelism is underscored by the fact that the two channels affect the same sound, which at times is done simultaneously through playing and physical gesture. It is straight-forward that playing in interactive music can produce and affect electronic sound, but it is interesting to note that the production and affect of sound through physical gestures is just as readily accepted as a valid connection in the audience's perception, once the initial surprise has ended.

Hand gestures affect the music procedurally generated in the piece *Stgo* for laptop computer (Graugaard 2005b). Laptop performance is mostly based on simple and effective manipulation of pre-recorded material and has developed its own performance practise bordering on the enigmatic (Cascone 2000). Informing laptop music with body gestures makes it possible to instill human expressiveness in real-time and establishes an appreciable performer-performance link (Graugaard 2006b). Significant data concerning the hand gestures determine perceptually significant elements of the procedurally generated music. The interface permits the easy transformation of the sound spectrum by hand gestures and the result is a dialogue where hand movements take on their meaning based on the present state of the music. The sensation of shaping the sound is strong and immediate and is readily perceived by the audience, even though not all relation hands-sound are directly traceable. This is certainly because the gestures and the resulting sound are intuitively related, but also very much because the hands are our primary shaping tool: We shape physical objects with them, and the imaginary shaping of sound is directly translated and readily felt.

The *SoundGove II* is a real time system for generating and affecting music and visuals through surface electromyograph (sEMG) muscular electrical activity sensing (Graugaard 2005a). The *SoundGlove II* is a non-utilitarian art installation. Its artistic purpose is to represent and intermingle a user's conscious and subconscious states, and to present it in an intuitively appreciable form. The *SoundGove II* does not require any skills or previous knowledge for its use, and is accesible to the ages from 10 years and up. It facilitates informal experience of artistic creation to non-expert users, an approach based on the assumption that all human beings are creative in their actions (Dalton 2004). For our purpose we make a query into the physiological signals the user emits to establish correlations between emotion state and bio signals, and audiovisual output. True emotion tracking with EMG data alone doesn't seem viable, and the *SoundGove II* suggests emotions when the sEMG's amplitude envelope as primary information source is used as a force or arousal indicator, and secondary sEMG data further classify the information presented to the system. This makes it possible to deduce relative emotional states, or directions of emotions.

5 DISCUSSION AND FUTURE WORK

Managing the digital analysis/synthesis step from the perspective of dynamic IRF is a powerfull means for contextualizing the digital component in interactive music (a subgroup of human-machine interaction). The prevailing model for interactive music only approaches the traditional music performance model because the digital component does not take the perceptual qualities of its own output into account. We are therefore investigating a symmetrical information loop (Fig. 5-1) which raises a range of issues concerning the infinite feedback loop which have been avoided with the traditional model.

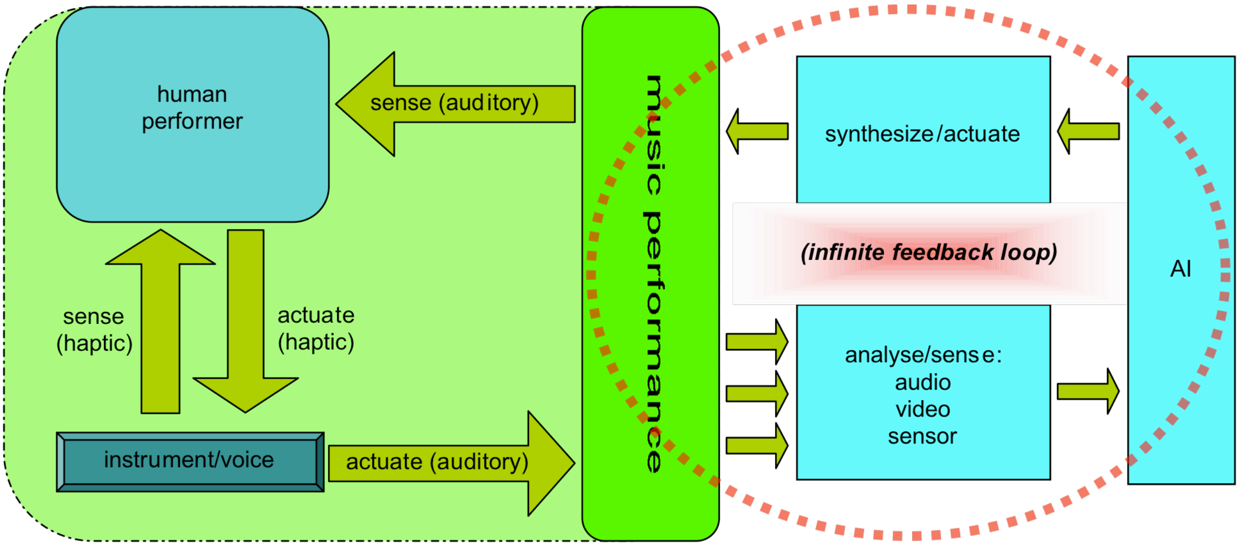

Fig. 5-1*. All-inclusive expression actuate-sense flow for interactive music.*

Music performance expresses emotions which relate to sensimotoric actions. Querying a user's emotion indicators provides more general information and such information are also present in part- and full-body movements and physiological signals. The interaction can therefore include other perceptual information channels, and the model is consequently applicable to the larger area of mutimedia systems. We are presently undertaking practical work with interactive systems exploring the all-inclusive expression actuate-sense model in a body of commissioned works for public presentation November 2006 – October 2008.

Interactive music and multimedia systems rely on an intuitive information channel for relevance assessment, where the user feedback is selective in that it contributes to shaping the immediate direction the system output takes. Insights into IRF in these areas may prove relevant to other, language based or online information retrieval systems where a user searches for information in an informal manner. However, a significant challenge will be the non-intrusive probing of such emotion data. At the moment of writing of this paper eye movement detection is one area that seems to hold such future possibilities.

6 REFERENCES

Bresin et al. 2000

Bresin, R., Friberg A.: Emotional Coloring of Computer-Controlled Music Performance; *Computer Music Journal* 24:4:44-63.

Camurri et al. 1999

Camurri, A., Hashimoto, S., Ricchetti, M., Suzuki, K., Trocca, R., Volpe, G.: *KANSEI Analysis of Movement in Dance/Music Interactive Systems*.

Camurri et al. 2000

Camurri, A., Dillon, R., Saron, A.: An Experiment on Analysis and Synthesis of Musical Expressivity. *Proceedings of the XIII Colloquium on Music Informatics*. L'Aquila, Italy.

Camurri et al. 2002

Camurri, A., Trocca, R., Volpe, G.: Interactive Systems Design: A KANSEI-based Approach. *Proceedings of NIME 2002*. Dublin, Ireland.

Canazza et al. 2001

Canazza, S., De Poli, G., Rodà, A., Vidolin, A., Zanon, P.: Kinematics-energy Space for Expressive Interaction in Music Performance. *Proceedings of MOSART, Workshop on current research directions in Computer Music*; pp. 35-40; Barcelona, Spain.

Canazza et al. 2004

Canazza, S.; De Poli, G.; Drioli, C.; Roda, A.; Vidolin, A.: Modeling and Control of Expressiveness in Music Performance. *Proceedings of the IEEE*, Volume 92, Issue 4, April 2004 pp. 686 - 701.

Cascone 2000

Cascone, K.: The Aesthetics of Failure: "Post-Digital" Tendencies in Contemporary Computer Music. *Computer Music Journal*, 24:4, Winter 2000, pp. 12-18. Massachusetts Institute of Technology.

Dahl et al. 2004

Dahl, S., Friberg, A.: Expressiveness of Musician's Body Movements in Performances on Marimba. In A. Camurri, G. Volpe (eds.): *Gesture-based Communication in Human – Computer Interaction*. LNAI 2915 (pp.479-486), Berlin, Springer Verlag.

Dalton 2004

Dalton, B.: Creativity, Habit, and the Social Products of Creative Action: Revising Joas, Incorporating Bourdieu. *Sociological Theory* 22:4, December 2004; pp. 603-622.

Friberg 2004

Friberg, A.: A fuzzy analyzer of emotional expression in music performance and body motion; in J. Sundberg & B. Brunson (Eds.): *Proceedings of Music and Music Science*, Stockholm, October 28-30 2004.

Friberg et al. 2002

Friberg, A., Schoonderwaldt, E., Juslin, P. N., Bresin, R.: Automatic Real-Time Extraction of Musical Expression. *Proceedings of the 2002 International Computer Music Conference*, pp. 365-367; Gothenburg, Sweden.

Gabrielsson et al. 1996

Gabrielsson, A., Juslin, P.: Emotional Expression in Music Performance. *Psychology of music* 24, pp. 68-91.

Graugaard 2004a

Graugaard, L.: Open and Closed Form in Interactive Music. In U. Kock Wiil (ed.) *Computer Music Modeling and Retrieval, Second International Symposium CMMR 2004*; Esbjerg, Denmark, May 26-29 2004; Revised Papers; pp. 149-157; Springer Verlag Berlin Heidelberg; ISBN 3-540-244858-1.

Graugaard 2004b

Graugaard, L.: *La Quintrala*. Opera for five singers and interactive, computer generated accompaniment. Premiered September 2nd 2004 at Den Anden Opera, Copenhagen, Denmark.

Graugaard 2005a

Graugaard, L.: The SoundGlove II: using sEMG data for intuitive audio and video affecting in real-time. *Proceedings of The Third Annual Conference in Computer Game Design and Technology*; pp. nn-nn; Liverpool, UK.

Graugaard 2005b

Graugaard, L.: *Stgo*. Premiered December 6th 2005 by Lars Graugaard; Cech Santiago, Chile.

Graugaard 2006a

Graugaard, L.: *Gestures You Made*. Composition for oboe and interactive, computer generated accompaniment. Premiered February 12th 2006 by Eydís Franzdottír; Dark Music Days 2006, Reykjavik, Iceland.

Graugaard 2006b

Graugaard, L.: Applying a Performer's Physical Gestures to Sound Synthesis in Real-time. *Proceedings of The International Computer Music Conference 2006*, New Orleans, Louisiana.

Graugaard 2006c

Graugaard, L.: *Gesture and Emotion in Interactive Music: Artistic and Technological Challenges*, PhD Dissertation, submitted March 2006, Music Department, Oxford Brookes University.

Jensen et al. 2006

Jensen, K., Graugaard, L.: Musical Expression with the D^2MMG Interface. *Proceedings of the 2nd ConGAS International Symposium on Gesture Interfaces for Multimedia Systems*, Leeds, England, May 9-10 2006.

Juslin 2000

Juslin, P. N.: Cue Utilization in Communication of Emotion in Music Performance: Relating Performance to Perception. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 1797-1813.

Leman 2006

Leman, M.: *Embodied Music Cognition and Mediation Technology*. In press.

Lesaffre et al. 2003

Lesaffre, M., Leman, M., Tanghe, K., De Baets, B., De Meyer, H., Martens, J.-P.: User-Dependent Taxonomy Of Musical Features As A Conceptual Framework For Musical Audio-Mining Technology. *Proceedings of the Stockholm Music Acoustics Conference*. Stockholm, Sweden.

Lippe 1996

Lippe, C.: A Look at Performer/Machine Interaction Using Real-Time Systems. *Proceedings of the International Computer Music Conference 1996*, pp. 116-117, Hong Kong.

Martin et al. 1998

Martin, K. D., Scheirer, E. D., and Vercoe, B. L.: Music Content Analysis through Models of Audition. *Proceedings of the ACM Multimedia '98 Workshop on Content Processing of Music for Multimedia Applications*, Bristol UK.

Moore 1990

Moore, F. R.: *Elements of Computer Music*. Prentice Hall, 1990, ISBN 0-13-252552-6.

Scherer et al. 2001

Scherer, K. R., Zentner, M. R.: Emotional Effects of Music: Production Rules. In Juslin, P.N. & Sloboda, J.A. (eds.): *Music and emotion: theory and research*. Oxford; New York, Oxford University Press.

Smalley 1997

Smalley, D.: Spectromorphology: Explaining Sound-shapes. *Organised Sound* 2(2), pp. 107-126; Cambridge University Press, United Kingdom 1997.

Sundberg 1993

Sundberg, J., D.: How Can Music Be Expressive? *Speech Communication*13:239–253.

Wanderley 2000

Wanderley, M.: Mapping Strategies. In Wanderly, M., Battier, M. (eds.): *Trends in Gestural Control of Music*; Ircam, Paris, 2000.

Zannos et al. 1997

Zannos, I., Modler, P., Naoi, K.: Gesture Controlled Music Performance in a Real-time Network. *Proceedings of KANSEI - The Technology of Emotion Workshop*, pp. 60–63, 1997.

Zanon et al. 2003

Zanon, P., & De Poli, G.: Estimation of Parameters in Rule Systems for Expressive Rendering of Musical Performance. *Computer Music Journal*, 27(1), 29-46.