

# I-Sounds

## Emotion-Based Music Generation for Virtual Environments

Ricardo Cruz<sup>1</sup>, António Brisson<sup>1</sup>, Ana Paiva<sup>1</sup>, and Eduardo Lopes<sup>2</sup>

<sup>1</sup> Instituto Superior Técnico  
Intelligent Agents and Synthetic Characters Group, INESC-ID, Lisboa, Portugal  
<sup>2</sup> Universidade de Évora, Music Department, Évora, Portugal  
{ricardo.cruz,antonio.brisson,ana.paiva}@gאים.inesc-id.pt  
el@uevora.pt

**Abstract.** I-Sounds aims to increase the Affective Bandwidth of an Interactive Drama system called I-Shadows, implementing a fully emergent system that generates affective music, based on musical theory and on the emotional state of the characters.

**Keywords:** Automatic music generation, sound, emotions, affective computing.

## 1 Music Generation Systems

We can identify two major approaches to music generation. Assisted composition uses pre-composed musical elements and/or user parameterization, such as Nakamura's [1] and Downie's [2] work. The alternate approach is the synthesis of composition algorithms based on musical knowledge and composition theory elements, the case of Herman [3,4]. Some relevant research focuses on music performance rather than music generation to convey emotions. Roberto Bresin's work [5] is a reference in this field.

## 2 The I-Sounds System

The I-Sounds system defines its own affective model and domains (e.g. a music score representation). Integration with applications is possible through the coding of an "I-Sounds driver". This mechanism assures the independence and generality of I-Sounds. By allowing programmers to extend I-Sounds models we intend to improve the system's flexibility. The system divides into three levels. The first includes the I-Sounds driver described above responsible for translating the input of the application into the I-Sounds domain. The second divides into an affective module which receives affective input from the driver and uses it to evolve an affective context and a composition module, an asynchronous music generation pipeline whose filters (stages) implement the composition heuristic, fed with information provided by the affective context of the previous module

transforming it into a music score. Finally the third level is responsible for the output of musical scores (e.g. MIDI sequences).

I-Sounds mapping heuristic will be mainly focused on rhythm. Lopes [6] develops in his PhD thesis a theory about rhythm and metre based on the two basic rhythmic qualities of pulse salience and kinesis. Pulse salience refers to the “emphasis” that a single pulse has on a rhythmic sequence, resulting from three components, the pulse’s metric position, agogic accentuation and rhythmic cell accentuation. Kinesis is related with motion induction. There is an inverse relation between kinetic potential and the metrical stability of a pulse. In what respects to melody, we will be using only the first five degrees of the diatonic scale, enough to use the diatonic major and minor modes closely related with happiness and sadness respectively.

In spite of being a general system able to integrate with different applications under different environments, I-Sounds will be tested in a specific context, more precisely in an interactive narrative environment called I-Shadows, [7] which is being developed in our research group at INESC-ID [7]. One of the goals of I-shadows is to provide an environment where children learn how to create stories and act them out in character in front of an audience, in an expressive way. We expect I-Sounds to enlarge the Affective Bandwidth between the System and the Users adding another Affective Channel to the interaction, Sound.

### 3 Future Work

I-Sounds is currently a work in progress. We think that using a tested musical theory will provide a solid scientific basis to the system. Future work may contemplate; the use of the whole scale degrees as well as other scales and modes, extended affective models with complex emotions and exploring I-Sounds output in different ways (e.g. putting a synthetic character dancing a generated sequence) rather than producing audible signals.

### References

1. Nakamura, J., Kaku, T., Hyun, K., Noma, T., Yoshida, S.: Automatic background music generation based on actors’ mood and motions. *he Journal of Visualization and Computer Animation* 25(5), 247–264 (1998)
2. Downie, M.: Behavior, animation, music: the music and movement of synthetic characters. Master’s thesis, Massachusetts Institute of Technology (2001)
3. Quincey, A.: Herman. Master’s thesis, University of Edinburgh (1998)
4. Stapleford, T.: The harmony, melody, and form of herman, a real-time music generation system. Master’s thesis, University of Endinburgh (1998)
5. Bresin, R.: Virtual Virtuosity, Studies in Atomatic Music Performance. PhD thesis, Kungliga Tekniska Högskolan, Stockholm (2000)
6. Lopes, E.: Just in Time: Towards a theory of rhythm and metre. PhD thesis, University of Southampton (2003)
7. Brisson, A., Fernandes, M., Paiva, A.: Children as affective designers. In: HUMAINE WP9 workshop, Kista, Sweden (2006)