ABSTRACT
How can we provide interfaces to synthesis algorithms that will allow us to manipulate timbre directly, using the same timbre-words that are used by human musicians to communicate about timbre? This paper describes ongoing work that uses machine learning methods (principally genetic algorithms and neural networks) to learn (1) to recognise timbral characteristics of sound and (2) to adjust timbral characteristics of existing synthesized sounds.

Keywords
timbre; natural language; neural networks

1. INTRODUCTION
When human musicians “interface” with other human musicians, we do so in words. One way in which we use words is to communicate about timbre: we say “can you make the sound shine more”, “I’d like to get a really gritty sound”, “let’s try to play that more warmly”.

This ability to use natural language descriptions of timbre (whether “straight” or metaphorical) is typically absent from interfaces with music technology devices. As a result, users have to either have a very strong understanding of the underlying mechanisms that produce the sound, or a large amount of “trial-and-error” experience with generating timbral changes within a system [5, 11]. A small number of attempts [2, 7] have attempted to create systems that offer an intuitive interface to timbre; however, there appears to be little recent work in this area [9].

By timbre we will mean the micro-level spectral characteristics of sound as discussed by Wishart [10], as opposed to the gross timbral distinctions [6] used e.g. in the MPEG-7 standard.

In this paper we discuss ongoing work which applies machine learning methods to associate changes in synthesis parameters with changes in timbre described by adverbs and adjectives. Overall the system is structured as illustrated in figure 1. Before the system is used, a training process is carried out whereby the system learns to associate features of sounds with certain timbre-description words that have been allocated by a human listener. Then,

Figure 1: Indirect modification of timbres.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
NIME 06, June 4-8, 2006, Paris, France
Copyright remains with the author(s).
Figure 2: Adjusting timbres with sliders.

We are currently pursuing a number of future directions for this work:

- Improved timbre recognition, for example by using ear-like pre-processing steps [4].
- Focusing more effort on having the system not adjust those characteristics of the sound that are not relevant to the current
- Interfacing this to e.g. MIDI control wheels to allow on-the-fly live manipulation of timbre.
- Creating systems that learn in advance the directions in parameter space which affect a particular timbre change, rather than running the parameter-adjustment algorithms on the fly. This may be a prerequisite for the use of the system in a live environment.

6. REFERENCES