SOUND MORPHING BY FEATURE INTERPOLATION

Marcelo Caetano

Sound morphing has been used in music compositions, in synthesizers, and even in psychoacoustic experiments, notably to study timbre spaces. When morphing musical instrument sounds, we usually want to obtain hybrid sounds that are *perceptually* intermediate between source and target across timbre dimensions, such that the intermediate sounds would correspond to hybrid instruments between source and target. A very challenging aspect of this problem is to control the transformation with a single parameter α, called morphing or interpolation factor. Ideally, we want the morphing factor α to control perceptually related features of the transformation, such that the morph should be perceptually halfway when $\alpha = 0.5$, for instance. Most morphing techniques proposed in the literature use the interpolation principle, which consists in interpolating the parameters of the model used to represent the sounds regardless of features. In this work, parameter refers to coefficients from which we can resynthesize sounds, while feature refers to coefficients used to describe or identify a particular aspect of a sound. Usually, we cannot resynthesize sounds directly from feature values. The basic idea behind the interpolation principle depicted is that if we can represent different sounds by simply adjusting the parameters of a model, we should obtain a somewhat smooth transition between two (or more) sounds by interpolating between these parameters These authors often conclude that the linear interpolation of the parameters of known models does not correspond to linearly varying perceptually relevant features. If we want the result to sound perceptually intermediate, we need to develop techniques to interpolate perceptually motivated features. For such, we adopt the morphing by feature interpolation principle, where the goal is to synthesize a hybrid sound whose feature values are intermediate between source and target. When the features capture perceptually relevant information, the morphed sound whose features are interpolated is perceptually intermediate. Ideally, we would like to be able to interpolate in the feature space and retrieve the set of parameters that correspond to the interpolated feature values. Unfortunately, this is a notoriously difficult problem to solve since most features commonly used do not allow direct inversion for resynthesis, particularly when the features are correlated to perceptual characteristics of sounds. The features we use are acoustic correlates of salient timbre dimensions derived from perceptual studies, so sounds whose feature values are intermediate between two would be placed between them in the underlyibg timbre space used as guide. In this work, we describe techniques to obtain morphed sounds whose values of features are as close as possible to the interpolated feature values. Ideally, we want the feature values to vary linearly when the morphing factor varies linearly. We measure the perceptual impact of the morphed sounds directly by the feature values, using them as an objective measure with which to evaluate the results. Thus we consider that the morphed sounds change perceptually linearly when the corresponding feature values vary linearly.