
Abstract

A method is described for estimating telephone handset nonlinearity by matching the spectral magnitude of the distorted signal to the output of a nonlinear channel model, driven by an undistorted reference. This "magnitude-only" representation allows the model to directly match unwanted speech formants that arise over nonlinear channels and that are a potential source of degradation in speaker and speech recognition algorithms. As such, the method is particularly suited to algorithms that use only spectral magnitude information. The distortion model consists of a memoryless polynomial nonlinearity sandwiched between two finite-length linear filters. Minimization of a mean-squared spectral magnitude error, with respect to model parameters, relies on iterative estimation via a gradient descent technique, using a Jacobian in the iterative correction term with gradients calculated by finite-element approximation. Initial work has demonstrated the algorithm's usefulness in speaker recognition over telephone channels by reducing mismatch between high- and low-quality handset conditions.