This invention proceeds from the consideration that existing artificial bandwidth extension schemes may result in a degradation to the overall perceived naturalness of the extended audio signal. This degradation may be especially prevalent for the overall perception of sibilant sounds.

Embodiments aim to address the above problem.

There is provided according to a first aspect of some embodiments a method comprising: generating an excitation signal from an audio signal, wherein in the audio signal comprises a plurality of frequency components; extracting a feature vector from the audio signal, wherein the feature vector comprises at least one frequency domain component feature and at least one time domain component feature; determining at least one spectral shape parameter from the feature vector, wherein the at least one spectral shape parameter corresponds to a sub band signal comprising frequency components which belong to a further plurality of frequency components; and generating the sub band signal by filtering the excitation signal through a filter bank and weighting the filtered excitation signal with the at least one spectral shape parameter, wherein the spectral shape parameter is a sub band energy level value and the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal.

According to an embodiment the method may when generating the excitation signal comprise generating a residual signal by filtering the audio signal with an inverse linear predictive filter; filtering the residual signal with a post filter stage comprising an auto regressive moving average filter based on the linear predictive filter; and generating the excitation signal by up sampling and spectrally folding the output from the post filter stage.

The post filter stage may further comprise a spectral tilt filter and a harmonic filter.

The frequency components of the sub band signal may be distributed according to a psychoacoustic scale comprising a plurality of overlapping bands, and the frequency characteristics of the filter bank may correspond to the distribution of frequency components of the sub band signal.

gradient at points in the audio signal which result in a change in direction of the waveform of the audio signal; a ratio of the energy of a frame of the audio signal to the energy of a previous frame of the audio signal; and a voice activity detector indicating whether a frame of the audio signal is classified as active or inactive.

The method may further comprise combining the sub band signal with the audio signal to provide a bandwidth extended audio signal.

According to a second aspect of some embodiments there is provided an apparatus comprising at least one processor and at least one memory including computer code, the at least one memory and the computer code configured to with the at least one processor cause the apparatus to at least perform: generating an excitation signal from an audio signal, wherein in the audio signal comprises a plurality of frequency components; extracting a feature vector from the audio signal, wherein the feature vector comprises at least one frequency domain component feature and at least one time domain component feature; determining at least one spectral shape parameter from the feature vector, wherein the at least one spectral shape parameter corresponds to a sub band signal comprising frequency components which belong to a further plurality of frequency components; and generating the sub band signal by filtering the excitation signal through a filter bank and weighting the filtered excitation signal with the at least one spectral shape parameter, wherein the spectral shape parameter is a sub band energy level value and the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal.

According to an embodiment the apparatus when the at least one memory and the computer code is configured to with the at least one processor to cause the apparatus to at least perform generating the excitation signal the apparatus may be further configured to perform; generating a residual signal by filtering the audio signal with an inverse linear predictive filter; filtering the residual signal with a post filter stage comprising an auto regressive moving average filter based on the linear predictive filter; and generating the excitation signal by up sampling and spectrally folding the output from the post filter stage.

of an overlapping band of the audio signal; a value representing a centroid of the frequency domain spectrum of the audio signal; and a value representing the degree of flatness of the frequency domain spectrum.

The at least one time domain component feature of the feature vector may comprise at least one of the following: a gradient index based on the sum of the gradient at points in the audio signal which result in a change in direction of the waveform of the audio signal; a ratio of the energy of a frame of the audio signal to the energy of a previous frame of the audio signal; and a voice activity detector indicating whether a frame of the audio signal is classified as active or inactive.

The at least one memory and the computer code is further configured to perform combining the sub band signal with the audio signal to provide a bandwidth extended audio signal.

There is according to a third aspect of some embodiments a computer program product in which software code is stored in a computer readable medium, wherein said code realizes the following when being executed by a processor: generating an excitation signal from an audio signal, wherein in the audio signal comprises a plurality of frequency components; extracting a feature vector from the audio signal, wherein the feature vector comprises at least one frequency domain component feature and at least one time domain component feature; determining at least one spectral shape parameter from the feature vector, wherein the at least one spectral shape parameter corresponds to a sub band signal comprising frequency components which belong to a further plurality of frequency components; and generating the sub band signal by filtering the excitation signal through a filter bank and weighting the filtered excitation signal with the at least one spectral shape parameter, wherein the spectral shape parameter is a sub band energy level value and the sub band energy level value is attenuated when the power of the audio signal approaches an estimate of the level of noise in the audio signal.

According to an embodiment the computer program product when the code realizes generating the excitation signal when being executed by a processor the code may further realize: generating a residual signal by filtering the audio signal