PROVISIONAL PATENT APPLICATION

Filed Under 35 U.S.C. §111(b)

METHOD AND APPARATUS FOR REAL-TIME HARMONIC ISOLATION AND SELECTIVE PROCESSING OF AUDIO SIGNALS

Inventor: Walter Karl Kueffer III

Residence: Winter Garden, FL 34787

Filing Date: ____________________

TITLE OF INVENTION

Method and Apparatus for Real-Time Harmonic Isolation and Selective Processing of Audio Signals

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority as a provisional patent application under 35 U.S.C. §111(b). No prior related applications are claimed.

ABSTRACT

[0002] A method and apparatus for real-time isolation and selective processing of harmonic components within an audio signal. The system performs pitch detection on an input audio signal to determine a fundamental frequency, then constructs a spectral mask in the frequency domain that selectively passes or attenuates individual harmonic partials based on a user-controlled parameter corresponding to a virtual node position along a vibrating string. The spectral mask is applied via Short-Time Fourier Transform (STFT) processing, preserving the original phase relationships and natural amplitude characteristics of the selected harmonics while suppressing undesired partials. The system further enables independent signal processing of isolated harmonic content and residual fundamental content through a dual-bus architecture, allowing different audio effects (such as distortion, reverb, delay, filtering, or spatial processing) to be applied selectively to harmonic versus fundamental components. The invention may be embodied as a hardware effects pedal with expression pedal control, a software audio plugin (VST/AU/AAX), or an embedded DSP system. The approach is distinguished from prior art pitch-shifting and octave-generation methods in that it extracts harmonics already present in the source signal rather than generating new pitched content.

BACKGROUND OF THE INVENTION

Field of the Invention

[0003] The present invention relates to audio signal processing, and more particularly to methods and apparatus for isolating and selectively processing harmonic partials within a musical audio signal in real time.

Description of Related Art

[0004] **Prior Art: Pitch Shifting.** Existing audio effects for manipulating harmonic content primarily rely on pitch shifting — the transposition of an entire audio signal to a different frequency. Products such as the Digitech Whammy, Boss PS-6 Harmonist, and Eventide PitchFactor use time-domain granular synthesis or phase vocoder methods to shift the pitch of the entire signal up or down by a specified interval. These devices generate a new signal at a different pitch; they do not isolate or extract harmonics already present in the source material.

[0005] **Prior Art: Octave Generation.** Octave generators such as the Electro-Harmonix HOG2 (Harmonic Octave Generator), TC Electronic Sub'N'Up, and EHX Pitch Fork generate additional copies of the input signal at octave intervals (one or two octaves above or below). These products add content to the signal through pitch-shifted duplication. They do not analyze or selectively filter the existing harmonic content of the input.

[0006] **Prior Art: Harmonic Enhancement.** Products such as the Boss MO-2 Multi Overtone use proprietary "Multi-Dimensional Processing" to add harmonically-related content to a guitar signal. These enhance or add harmonic content but do not isolate or separate the existing harmonic components from the fundamental.

[0007] **Prior Art: Source Separation.** Academic techniques including Harmonic/Percussive Source Separation (HPSS, Fitzgerald 2010) and Sinusoidal Modeling Synthesis (SMS, Serra & Smith 1990) provide frameworks for decomposing audio signals into harmonic and non-harmonic components. However, these techniques have not been applied to the specific problem of simulating string harmonic node positions in real time for musical performance, nor have they been combined with dual-bus selective processing architectures for independent effects application to harmonic versus fundamental components.

[0008] **The Gap in the Art.** When a guitarist performs a natural harmonic by lightly touching a vibrating string at a node point, specific standing wave modes are physically suppressed. Only harmonics whose standing wave patterns have a natural node at the touch point survive. This creates the characteristic bell-like, crystalline tone of natural and artificial harmonics. No prior art system simulates this physical phenomenon electronically by analyzing the harmonic content of an audio signal and selectively passing only those harmonics that would survive at a specified node position. Furthermore, no prior art system separates the harmonic content from the fundamental content for independent effects processing.

SUMMARY OF THE INVENTION

[0009] The present invention provides a method, apparatus, and computer-readable medium for isolating harmonic partials from an audio signal in real time, based on a user-selectable virtual node position.

[0010] In one embodiment, the method comprises: (1) Receiving an audio input signal containing a musical note with a fundamental frequency and associated harmonic partials; (2) Detecting the fundamental frequency (f_0) of the audio signal using autocorrelation-based pitch detection; (3) Computing a Short-Time Fourier Transform (STFT) of the audio signal; (4) Constructing a spectral mask in the frequency domain that assigns gain values to frequency bins corresponding to harmonic partials of the detected fundamental, where the gain values are determined by a virtual node position parameter; (5) Applying the spectral mask to the STFT representation; (6) Reconstructing the processed audio signal via inverse STFT with overlap-add synthesis; (7) Optionally mixing the processed signal with the original dry signal according to a wet/dry parameter.

[0011] In a further embodiment, the system separates the audio signal into a harmonic bus (containing the masked harmonic content) and a fundamental bus (containing the residual), enabling independent effects processing of each component before recombination.

[0012] The virtual node position parameter may be controlled by an expression pedal, a knob, a MIDI controller, an automation lane in a digital audio workstation, or any other continuous control source.

DETAILED DESCRIPTION OF THE INVENTION

System Architecture

[0013] Referring to Figure 1 (described below), the system comprises the following signal processing stages connected in series.

[0014] **Stage 1: Audio Input.** The audio input is received from a musical instrument (guitar, bass, violin, piano, voice, or other pitched source) via an analog-to-digital converter operating at a sample rate of at least 44,100 Hz with at least 16-bit resolution. The input is assumed to be monophonic (single pitch at a time), though the system may process polyphonic signals with reduced accuracy.

[0015] **Stage 2: Pitch Detection.** The fundamental frequency (f_0) of the input audio is estimated using autocorrelation-based pitch detection. For each analysis frame: (1) The frame is windowed using a Hann window of length W samples (preferably W = 2048 to 4096 samples). (2) The autocorrelation function is computed via FFT: $R(\tau) = IFFT(|FFT(x \cdot w)|^2)$. (3) The autocorrelation is normalized: $R'(\tau) = R(\tau) / R(0)$. (4) The peak of $R'(\tau)$ is located within the range corresponding to expected musical pitches (e.g., 55 Hz to 1000 Hz), yielding a lag estimate. (5) Parabolic interpolation refines the lag estimate to sub-sample precision. (6) The confidence of the pitch estimate is taken as the normalized autocorrelation value at the peak. (7) Pitch estimates with confidence below a threshold (e.g., 0.25) are rejected, and the previous valid estimate is retained. (8) Temporal smoothing is applied to prevent pitch jumps: if the new estimate differs from the previous by more than 15%, a weighted average is used (e.g., 0.7 × previous + 0.3 × new).

[0016] **Stage 3: Spectral Mask Construction.** For each STFT frame, a spectral mask M[k] is constructed for k = 0, 1, ..., N/2, where N is the FFT size. The mask determines the gain applied to each frequency bin. The mask construction proceeds as follows: (1) Initialize all mask values to a floor value (e.g., 0.003, representing approximately -50 dB suppression). (2)

Determine the virtual node position from the user control parameter P (0.0 to 1.0). (3) Map P to a harmonic divisor D through a piecewise linear function: P ∈ [0.0, 0.05]: D = 1 (no filtering, all harmonics pass); P ∈ [0.05, 0.2]: D interpolates from 1 to 2; P ∈ [0.2, 0.4]: D interpolates from 2 to 3; P ∈ [0.4, 0.6]: D interpolates from 3 to 4; P ∈ [0.6, 0.8]: D interpolates from 4 to 5; P ∈ [0.8, 1.0]: D interpolates from 5 to 6. (4) For each harmonic number h = 1, 2, ..., H_max: compute the harmonic frequency f_h = h × f_0; determine the weight for this harmonic from the bracketing node positions using linear interpolation; for integer divisor D_a: weight_a = 1.0 if h mod D_a = 0, else floor_value; for integer divisor D_b: weight_b = 1.0 if h mod D_b = 0, else floor_value; final weight = (1 - t) × weight_a + t × weight_b, where t is the interpolation factor; identify the center frequency bin: k_c = round(f_h / (sr / N)); apply the weight to bins within a window around k_c, with flat-top response and tapered edges; the window half-width is set to approximately 35% of the bin spacing between adjacent harmonics to prevent overlap.

[0017] The resulting mask has values near 1.0 at frequencies corresponding to surviving harmonics and near 0.003 at all other frequencies.

[0018] Stage 4: Mask Application and Reconstruction. (1) The spectral mask is temporally smoothed with the previous frame's mask using an exponential moving average (e.g., α = 0.4 for the new mask, 1-α = 0.6 for the previous) to prevent audible discontinuities. (2) The mask is applied to the STFT frame by element-wise multiplication: Y[k] = X[k] × M[k]. (3) The inverse FFT reconstructs the time-domain frame. (4) Overlap-add synthesis with a synthesis window (Hann) reconstructs the continuous output signal. (5) The overlap-add normalization factor is computed and applied to ensure unity gain in steady state.

[0019] Stage 5: Wet/Dry Mixing. The processed (wet) signal is mixed with the original (dry) signal according to a user-controlled wet/dry parameter: output = wet_mix × processed + (1 - wet_mix) × dry.

[0020] Stage 6: Dual-Bus Processing (Extended Embodiment). In an extended embodiment, the system further comprises: (1) A harmonics bus carrying the output of Stage 4 (the masked harmonic content). (2) A fundamental bus obtained by subtracting the harmonics bus from the original signal, or by applying the inverse mask (1 - M[k]) to the STFT representation. (3) Independent effects processing chains applied to each bus (e.g., distortion, reverb, delay, filtering, compression, tremolo, ring modulation, bit reduction, spatial panning, or any combination thereof). (4) A mixing stage that recombines the independently processed buses. This dual-bus architecture enables an entirely new class of audio effects where different processing is applied to harmonic content versus fundamental content.

Physical Basis

[0021] The spectral mask simulates the physical phenomenon of touching a vibrating string at a node point. When a string is touched at position L/n (where L is the string length and n is an integer), only standing wave modes with harmonic numbers that are multiples of n can have a node at that position. All other modes have non-zero displacement at the touch point and are damped.

[0022] The mapping of the user control parameter P to the harmonic divisor D corresponds to touching the string at: D = 2: midpoint (12th fret on guitar) — keeps even harmonics (2, 4, 6, 8...); D = 3: one-third point (7th fret) — keeps multiples of 3 (3, 6, 9, 12...); D = 4: one-quarter point (5th fret) — keeps multiples of 4 (4, 8, 12...); D = 5: one-fifth point (approximately 4th fret) — keeps multiples of 5 (5, 10, 15...); D = 6: one-sixth point (approximately 3rd fret) — keeps multiples of 6 (6, 12, 18...).

BRIEF DESCRIPTION OF THE DRAWINGS

[0023] **Figure 1** is a block diagram showing the signal flow of the system from audio input through pitch detection, STFT analysis, spectral mask construction and application, ISTFT/overlap-add reconstruction, and wet/dry mixing to produce the final output. An expression pedal input feeds the mask construction module.

[0024] **Figure 2** shows three panels of spectral mask examples for a fundamental frequency of 110 Hz. Panel A shows the Node 2 mask (D=2) with peaks at 220 Hz, 440 Hz, 660 Hz, and 880 Hz (even harmonics) and suppression at 110 Hz, 330 Hz, 550 Hz, and 770 Hz. Panel B shows the Node 3 mask (D=3) with peaks at 330 Hz, 660 Hz, and 990 Hz (multiples of 3) and suppression at other harmonics. Panel C shows the Node 4 mask (D=4) with peaks at 440 Hz and 880 Hz (multiples of 4).

[0025] **Figure 3** shows two spectrogram panels (frequency versus time). The left panel shows an original guitar signal (A2, 110 Hz) with all harmonic partials visible as horizontal lines. The right panel shows the processed signal (Node 2) with only even harmonics remaining and odd harmonics clearly suppressed.

[0026] **Figure 4** is a spectrogram showing a sustained note being processed while the node position parameter sweeps from 0 (all harmonics present) through Node 2, Node 3, Node 4, and Node 5, with harmonics progressively disappearing as higher node positions are selected.

[0027] **Figure 5** is a block diagram showing the dual-bus architecture with the audio input splitting into a harmonics bus (through the spectral mask) and a fundamental bus (through the inverse mask), each feeding into independent effects chains that converge at a mixing stage.

[0028] **Figure 6** is an illustration of a vibrating string showing standing wave patterns for harmonics 1 through 6, with node points marked, and arrows indicating the correspondence between touching at L/2, L/3, L/4, etc. and the spectral mask divisor values D=2, 3, 4, etc.

CLAIMS

What is claimed is:

1. A method for processing an audio signal, comprising:

(a) receiving an audio input signal containing one or more harmonic partials of a fundamental frequency;

(b) detecting the fundamental frequency of the audio input signal;

(c) computing a frequency-domain representation of the audio input signal;

(d) constructing a spectral mask based on the detected fundamental frequency and a user-controlled node position parameter, wherein the spectral mask assigns a first gain value to frequency bins corresponding to harmonic partials that are integer multiples of a harmonic divisor determined by the node position parameter, and a second, lower gain value to other frequency bins;

(e) applying the spectral mask to the frequency-domain representation;

(f) reconstructing a time-domain output signal from the masked frequency-domain representation.

2. An audio effects apparatus comprising:

(a) an audio input for receiving an audio signal from a musical instrument;

(b) an analog-to-digital converter;

(c) a digital signal processor configured to perform pitch detection, spectral analysis, spectral mask construction, mask application, and signal reconstruction as described in claim 1;

(d) a digital-to-analog converter;

(e) an audio output;

(f) a user control interface comprising at least one of: a knob, an expression pedal input, or a footswitch;

wherein the user control interface controls the node position parameter of the spectral mask.

3. A non-transitory computer-readable medium storing instructions that, when executed by a processor, cause the processor to perform the method of claim 1, the instructions being packaged as an audio plugin conforming to at least one of the VST3, Audio Unit, or AAX plugin interface standards.

4. A method for selective audio processing, comprising:

(a) performing the method of claim 1 to produce a harmonics signal;

(b) computing a fundamental signal by subtracting the harmonics signal from the original audio input signal, or by applying an inverse spectral mask;

(c) independently processing the harmonics signal through a first audio effects chain;

(d) independently processing the fundamental signal through a second audio effects chain;

(e) combining the independently processed harmonics signal and fundamental signal to produce a final output signal;

wherein the first and second audio effects chains apply different audio processing operations.

5. The method of claim 1, wherein detecting the fundamental frequency comprises computing an autocorrelation function of a windowed frame of the audio input signal and locating a peak in the autocorrelation function within a range corresponding to musical pitches.

6. The method of claim 1, wherein the frequency-domain representation is a Short-Time Fourier Transform (STFT) and the reconstruction employs overlap-add synthesis.

7. The method of claim 1, wherein the node position parameter is continuously variable and the harmonic divisor is interpolated between integer values, such that the spectral mask transitions smoothly between discrete node positions.

8. The method of claim 1, wherein the spectral mask is temporally smoothed between consecutive analysis frames using an exponential moving average to reduce audible artifacts.

9. The method of claim 1, further comprising mixing the reconstructed output signal with the original audio input signal according to a wet/dry parameter.

10. The method of claim 1, wherein the spectral mask assigns gain values to frequency bins within a window around each harmonic frequency, the window having a flat-top region and tapered edges, and the window width being less than 50% of the bin spacing between adjacent harmonics.

11. The apparatus of claim 2, wherein the digital signal processor is an ARM Cortex-M7 processor or equivalent, and the apparatus is housed in a guitar effects pedal enclosure.

12. The apparatus of claim 2, further comprising an expression pedal input that controls the node position parameter continuously from a minimum to a maximum value.

13. The method of claim 4, wherein the first audio effects chain comprises distortion and the second audio effects chain comprises no distortion, such that distortion is applied to harmonic content but not fundamental content.

14. The method of claim 4, wherein the first audio effects chain comprises a reverberation effect and the second audio effects chain is a pass-through, such that reverberation is applied to harmonic content but not fundamental content.

15. The method of claim 4, wherein the harmonics signal is routed to a first stereo channel and the fundamental signal is routed to a second stereo channel, providing stereo separation of harmonic and fundamental content.

16. The method of claim 1, wherein the audio input signal is produced by any of: an electric guitar, an acoustic guitar, a bass guitar, a bowed string instrument, a keyboard instrument, a brass instrument, a woodwind instrument, or a human voice.

17. The method of claim 4, wherein the first and second audio effects chains are user-configurable, and the types of audio processing operations include any combination of: distortion, reverb, delay, compression, filtering, tremolo, ring modulation, bit depth reduction, pitch shifting, and spatial panning.

18. The method of claim 1, wherein the node position parameter is controlled by an envelope follower tracking the amplitude of the audio input signal, such that the harmonic isolation responds dynamically to playing intensity.

ABSTRACT OF THE DISCLOSURE

A method and apparatus for real-time isolation and selective processing of harmonic components within an audio signal. The system performs pitch detection on an input audio signal to determine a fundamental frequency, then constructs a spectral mask in the frequency domain that selectively passes or attenuates individual harmonic partials based on a user-controlled parameter corresponding to a virtual node position along a vibrating string. The spectral mask is applied via Short-Time Fourier Transform (STFT) processing. The system further enables independent signal processing of isolated harmonic content and residual fundamental content through a dual-bus architecture. The invention may be embodied as a hardware effects pedal, a software audio plugin, or an embedded DSP system.

REFERENCES CITED

U.S. Patent Documents

1. U.S. Patent 4,117,413 — "Musical effects device" (octave divider)

2. U.S. Patent 5,536,902 — "Method of pitch detection" (autocorrelation)

3. U.S. Patent 6,448,487 — "Pitch shifting" (time-domain)

Other References

4. Serra, X. & Smith, J.O. (1990). "Spectral Modeling Synthesis: A Sound Analysis/Synthesis System Based on a Deterministic plus Stochastic Decomposition." *Computer Music Journal*.

5. Fitzgerald, D. (2010). "Harmonic/Percussive Separation using Median Filtering." *13th International Conference on Digital Audio Effects (DAFx-10)*.

6. De Cheveigné, A. & Kawahara, H. (2002). "YIN, a fundamental frequency estimator for speech and music." *Journal of the Acoustical Society of America*.