

Signal Processing Overview
By Steven M. Miller

I) Signal Processing Overview

A) Basic (generic) signal processing types

- 1) Timbral (equalization)
 - a) timbre
 - b) texture
- 2) Dynamic (compression, limiting, expansion)
 - a) loudness
 - b) envelope
- 3) Spatial (reverb, panning, delay)
 - a) mental visual image
 - b) horizontal and vertical positioning

B) Interdependency

- 1) EQ
 - a) affects perceived loudness (dynamic)
 - b) affects perception of distance (spatial)
- 2) Dynamics
 - a) affects 'fullness' of sound (timbral)
 - b) can emphasize or de-emphasize attack and/or decay (timbral)
- 3) Spatial imaging
 - a) affects perception of 'clarity' (timbral)
 - b) can influence relative dynamics

II) Uses/abuses of signal processing

A) Corrective/remedial

- 1) system deficiencies
 - a) microphone frequency response
 - b) noise/hum in electronics
 - c) limited dynamic range
- 2) acoustical deficiencies in instrument or environment
 - a) poor instrument
 - b) overly dead or live room
- 3) operator error
 - a) poor microphone placement
 - b) improper level setting (gain structure)

B) Creative

- 1) artistic/musical choice
 - a) altered or unnatural timbre
 - b) exaggerated or animated spatial imaging
- 2) blending or contrasting sounds in a mix
 - a) accentuate different bands in audio spectrum
 - b) tonal and dynamic 'masking'

III) Signal processing particularities

A) EQ

- 1) Audible spectrum 20-20,000 Hz (waveform cycles-per-second) = 10 octaves
 - a) 5 arbitrary (idealized) bands
 - 1) Bass - 20-160 Hz - 3 octaves
 - a) very little (natural) musical energy below 60Hz
 - b) sound is felt as well as heard
 - c) adds sense of power
 - d) reduce mechanical noise, rumble
 - e) most perceptible adjustments 60-90Hz
 - 2) Low mids - 160-320Hz - 1 octave
 - a) too much = muddy or boomy
 - b) too little = thin or hollow
 - c) 200Hz = warmth
 - 3) Mids - 320-2560Hz - 3 octaves
 - a) ear's most sensitive range
 - b) dramatic effect
 - c) too much = listener fatigue
 - d) most musical fundamentals 200-1000

Signal Processing Overview
By Steven M. Miller

- 4) Upper mids - 2560-5120Hz - 1 octave
 - a) presence range
 - b) clarity of speech
 - c) too much = tinny or harsh
 - d) too little = indistinct or muffled
- 5) High end - 5120-20,000Hz - 2 octaves
 - a) brilliance range (cymbals, brass, strings)
 - b) consists primarily of musical overtones
 - c) vocal sibilance - 7,500-10,000Hz
 - d) 5120-12,000Hz = perceived as high frequencies
 - e) very little musical material above 12,000Hz
 - f) high frequency noise filter

B) Compression & Limiting

1) Definition

- a) dynamic range
 - 1) difference in level between softest and loudest sound
 - a) music - 120 dB
 - b) analog tape recorder (no noise reduction) - 65 dB
 - c) digital recorder - 90-100 dB
 - d) FM radio - 60dB
 - e) LP record - 60dB
 - f) CD - 90dB
- b) non-linear amplifier
 - 1) for given increase in input level, output is not always proportional 1:1
 - a) compressor: ratio of output level to input level decreases as input increases above threshold
 - b) limiter: constant output level regardless of increase in input level above threshold
 - c) expander: decreases output gain as input falls below threshold or increases output gain as input rises above threshold
- c) threshold
 - 1) point at which input signal level 'triggers' non-linear response
- d) ratio or slope
 - 1) proportion of change in input level to change in output level
 - 2) soft slope (1.5:1 to 2:1) preserves most of dynamics, evens out sound
 - 3) tighter slope (3:1 to 5:1) tightens up sound
 - 4) ratios above 10:1 are considered as limiting
- e) attack
 - 1) speed with which gain reduction takes place after signal rises above threshold
- f) release
 - 1) speed with which gain reduction returns to normal after signal drops below threshold
- g) gain reduction
 - 1) actual level difference in dB between processed and unprocessed signal

2) Uses

- a) Limiting
 - 1) matching dynamic ranges
 - a) control transients so you can record at high level w/o overload
- b) Compression
 - 1) smoothing - decrease difference between average and peak levels
 - a) voice narration over music
 - b) vocal fluctuations over music
 - c) changes in playing technique (pizz. vs. arco)
 - d) changes in mic-to-source distance

Signal Processing Overview
By Steven M. Miller

- 2) tonal balance
 - a) high or low frequencies often have different dynamics
 - 1) bass guitar strings not all equal volume
 - 2) brass high notes naturally louder (technique)
 - 3) smooth high-frequency spikes in voice - full sound
 - 3) ducking (voice-over compression)
 - a) voice actuates compression on music track, reduces level automatically
 - c) Expansion
 - 1) upward expansion
 - a) increase dynamic range of weak signals
 - 2) downward expansion
 - a) noise gate
 - 1) set threshold just above noise level
- C) Reverb, panning, delay
- 1) Definition
 - a) reverb
 - 1) repeated reflections of a sound wave after the sound source has ceased vibrating
 - 2) the perception of closely spaced and random multiple echoes reflected from one boundary to another in a determined space
 - b) panning
 - 1) changing the loudness level of a sound between two or more channels or speakers
 - c) delay
 - 1) (echo) the perceptible repetition of sound
 - 2) the time interval between a sound and each of its repeats
 - 2) Particularities
 - a) reverb
 - 1) direct sound
 - a) reaches the ear directly from the sound source
 - b) determines perception of sound source location, size and true timbre
 - 2) reflected sounds
 - a) general
 - 1) waves that bounce off surrounding surfaces before reaching the ear
 - 2) travel further, therefore arrive after the direct sound
 - 3) not absorbed equally at all frequencies
 - 4) timbre is coloured by characteristics of reflective surfaces
 - 5) ratio of level of direct to reflected sound determines perception of distance from sound source
 - b) early reflections
 - 1) reach the ear after reflecting off boundary and object surfaces
 - 2) echos which reach the ear within 50 msec of the original sound
 - 3) not usually perceived as discrete sound events
 - 4) within 40 msec-reflected sound is fused with original sound
 - c) RT60 - reverb/decay time
 - 1) the time it takes for the persisting sound to decrease to 60dB below its original level
 - 2) RT60 (length) main aural cue to size of room

Signal Processing Overview
By Steven M. Miller

- b) panning
 - 1) general
 - a) interaural intensity difference
 - b) most widely used method of sound positioning
 - c) not as effective other methods (delay, phase shift, stereo mic techniques)
 - d) not hardware intensive
 - e) traditional
 - f) mono-compatible
- c) delay
 - 1) general
 - a) interaural arrival-time difference
 - b) more effective than simple panning
 - c) hardware intensive
 - d) not mono compatible
 - 2) early delays
 - a) <40 msec perceived as part of original sound - fused
 - b) level - important cues to perception of location relative to sound source
 - c) length - important cues to perception of size of room
 - 3) discrete echo
 - a) >40 msec perceived as discrete sound event (echo) dependent on material
 - b) can be used to simulate outdoor echo effects