FH MMA SALZBURG – SOUND SYNTHESIS

SOUND SYNTHESIS 02

7. SAMPLERS AND SAMPLING TECHNOLOGY

A BIT OF HISTORY – BEFORE SAMPLING

PIPE ORGANS

Figure 1: The Arp-Schnitger organ in the Ludgerikirche, Norden (Germany)

The idea to imitate the sound of musical instruments with another instrument is not new: already centuries ago **pipe organs** included several registers that were at least inspired by other instruments and that sometimes really sounded very close to the originals (flutes, trumpets, cornets, bassoons, etc.).

To create a vast amount of different sound colors, the pipes were manufactured using different materials (wood, tin metals, etc.) and shapes (cylinder, cone, prism, etc.). For certain registers (such as trumpets, bassoon, cornets etc.) reeds were also used. Other registers used more than one row of pipes, slightly detuned against each other (*imitatio violistica*, *vox humana*) to obtain a chorus-like effect.

Typical organ register combinations were obtained through an *additive process*, adding to the original tone registers tuned in octaves and fifths (*Mixtur*, or *Organo Pleno*), or combinations of fifths and thirds (*Sesquialtera*). As the additional rows of pipes were tuned according to the natural overtones of the original tone, they tended to blend together creating a new timbre and were not perceived as separate notes.

This made the pipe organ one of the most powerful and flexible instruments of all times. However, it never really became very good at imitating orchestral string sounds, and it was not good at all at imitating percussive ones. In fact, the only way it could sound like a glockenspiel, was to add a glockenspiel inside the organ case (which has also been done in some large instruments, as well as adding bells and other effects).

Despite the possibility of emulating some real instruments, the organ had a wide spectrum of sounds that were characteristic for this instrument only.

HAMMOND ORGAN

Figure 2: Hammond B-3 Organ with Leslie Cabinet (ca. 1955, left) and register sliders (right)

Around 1930 the **Hammond** organ was introduced, an electro-mechanic music instrument using rotating tone-wheels for the sound generation. Initially intended as a replacement of theater organs, the Hammond used an additive sound process: the different timbres were obtained by adding together several rows of registers (each one sounding almost like a perfect sine wave) in combinations similar to the pipe organ mixtures (in intervals of octaves, fifths and thirds in relation to the fundamental tone).

The Hammond could not emulate any natural instrument except the organ – actually it was not even very good at sounding like a real organ: a single pipe organ register does not sound very much like a sine wave, except perhaps some flute registers. However, the Hammond had a very characteristic sound, so it became a very successful instrument on its own, especially in jazz and pop/rock music.

To make the relatively static “raw” sound of the Hammond more lively and interesting, the “Leslie” was introduced: this was a set of rotating loudspeakers (separate for highs and lows) that created a chorus-like effect, due to the changes in pitch caused by the Doppler-effect, and changes in colors caused by the reflections against the back and side walls. The speed of the rotation could be controlled by pedal, and therefore used for expressive purposes.

MELLOTRON

The next attempt to emulate real instruments was the **Chamberlin** (invented and developed by Harry Chamberlin in 1949-56), which eventually evolved into the **Mellotron**, first introduced in 1963. Really successful and reliably working models were first manufactured at the end of the '60s.

The Mellotron features under each key a magnetic tape loop and playback heads like in a tape recorder. When a key is pressed, the tape rotates, and the sound is read from the playback head. Each key requires a different tape recording at the right pitch. To change a register set, it is necessary to change all tapes manually with a different set!

Figure 3: Mellotron Mk VII and internal tape-based playback system

Because of the tapes being played back so many times, the sound quality progressively degraded with time, introducing frequency losses, distortion, noise etc. Furthermore, the transport mechanism was not totally reliable, causing wow & flutter (speed fluctuations).

Despite these limitations, this instrument was extremely successful and there was hardly any large rock group that didn't feature a Mellotron on stage in the '70s.

Both the Hammond organ and the Mellotron can be well heard – among other instruments – in the early productions of "Yes", for example in "Heart of The Sunrise" or "Close to the Edge".

HARDWARE DIGITAL SAMPLERS

It was not until analogue tapes could be replaced by digital sample recording and playback, that a relatively faithful reproduction of acoustic instruments became possible. The main advantage of storing the sound digitally was that unlimited reproductions of the sample were finally possible, without progressive degradation of the sound quality like in a Mellotron.

Between the most famous (and expensive) are the **Fairlight CMI** (Computer Music Instrument), introduced in 1979, and the **NED Synclavier** (New England Digital) introduced in 1977/78.

These instruments were very expensive (Fairlight CMI: 30 000 – 200 000 USD; NED Synclavier: 75 000 – 500 000 USD) and could only be afforded by top-tier artists and producers, and major recording studios.

NED SYNCLAVIER

Figure 4: NED Synclavier II System 9600

The specifications of the Synclavier II were incredible for the time, and even by modern standards: 16-bit stereo sampling, 100 kHz sampling frequency, up to 64 voices of polyphony, storage on hard-disk, multitrack recording capability, FM, additive and re-synthesis options, 32-768 MB RAM, external hard disk, etc.

The Synclavier is still considered as one of the most advanced electronic instruments ever built – in fact it is not only still being used by many studios and musicians, but it is also being supported: the software is continuously being improved, it is still possible to find replacement parts and complete refurbished instruments.
More info: http://www.500sound.com – http://www.synclavier.com/

The most accurate software reproduction of the Synclavier is a joint development between Arturia and Synclavier Digital. http://www.synclavier.com/synclavierv/

FAIRLIGHT CMI

Figure 5: Fairlight CMI IIx

The Fairlight CMI offered in the first versions “only” 8-bit sampling (later 12- and 16-bit) with limited sampled rate, but it used a hybrid digital/analogue signal path including analogue VCFs and VCAs. In spite of the low sound quality it became a “classic” and it is featured in countless albums throughout the ‘80s (for example in Peter Gabriel IV – 1982, Jean Michel Jarre “Zoolook” – 1984)

EMU EMULATOR, AKAI S-SERIES, ROLAND S-SERIES, ENSONIQ MIRAGE …

Figure 6: Among the first affordable samplers: EMU Emulator II and AKAI S900

In the mid '80s the first affordable samplers appeared: the **EMU Emulator** series, the **Ensoniq Mirage**, the **AKAI S900 and S1000** series, the **Roland S50**, etc. Some of these instruments had no filters at all and offered just 8 or 12-bit quantization, but still they were a huge success and were the first instruments that could be afforded by almost every musician.

The user interface varied greatly: while the S50 could be connected directly to a computer monitor, offering very advanced graphical editing capabilities similar to the Fairlight, the Ensoniq Mirage became famous as one of the most user-unfriendly instruments ever created: all sample editing, looping, multisampling, programming of envelopes etc. had to be done using a very small LED display and a numeric keypad – a real programming horror.

Figure 7: Roland S50 directly connected to VGA monitor

The **Roland S50** sampling keyboard (and the rackmount S550) had revolutionary features for the time: not only could the instrument be connected directly to a VGA monitor, allowing sample waveform visualization and editing, graphical display of envelopes and other parameters, etc. (features usually reserved to much more expensive models like the Fairlight CMI and the Synclavier): It also offered a pen/tablet combo as a controller (similar to modern Wacom pen/tablets) as well as a remote fader unit.

AKAI S 1000

In 1988 AKAI introduced the **S1000** sampler series, which featured for the first time 16-bit / 44.1 kHz linear PCM sampling in an affordable instrument, digital filters, internal effect sends, and up to 32 MB of RAM. This became one of the most successful samplers of all times. Other instruments, like the EMU Emulator III and the Synclavier, offered also 16-bit sampling, but at a much higher price tag (Emulator III: about 17.000 USD; Synclavier: see above).

Figure 8: AKAI S1000

Here an overview of the most popular samplers in the ’80s and ’90s:

- **EMU Emulator** (1981): 8-bit, 27,7 kHz, 128 KB RAM, 2-8 voices
- **EMU Emulator II** (1984): 8-bit, 27,7 kHz, up to 1 MB RAM, 8 voices, analog 4-pole LP resonant filters
- **Ensoniq Mirage** (1984): 8-bit, 128 KB of RAM, 8 voices, analog filters
- **KORG DSS-1** (1986): 12-bit / 48 kHz, up to 2 MB RAM, 8 voices, with analogue filters and envelopes
- **Roland S-50** (1986): 12-bit / 30 kHz, 750 KB of RAM (S-550: 1,5 MB), 16 voices, LP/HP digital filters
- **EMU Emulator III** (1987): 16-bit / 44.1 kHz stereo, 8 MB RAM, 16 voices
- **AKAI S900** (1986): 12-bit / 40 kHz, up to 750 KB of RAM, 8 voices
- **AKAI S1000** (1988): 16-bit / 44.1 kHz, up to 2-32 MB RAM, 16 voices, digital filters, internal effect sends
- **Roland S-770** (1989): 16-bit / 48 kHz, 24 voices, 2-18 MB RAM, digital filters and envelopes
- **Ensoniq ASR10** (1992): 16-bit /44.1 kHz, 31 voices, 2-16 MB RAM, LP&HP digital filters

Since the ’90s the standard for sampling has been 44.1 kHz and 16-bit (later also 24-bit), allowing a sampling fidelity comparable to CD audio. Among the most successful samplers of the ’90s are the **AKAI S3000** (1992), the **Roland S770** series, the **Emulator IV** and the **Ensoniq ASR10**.

Figure 9: Ensoniq ASR-10

This new generation of full digital samplers featured digital filters with resonance. While not quite comparable in quality with true analogue filters (like those used in the first hybrid sampler generation), or with the sort of advanced digital filters available today in software samplers, they were already a great improvement over the first generation of digital filters (like those featured in the AKAI S1000 or Roland S50).

SOFTWARE SAMPLERS

NEMESYS / TASCAM GIGASTUDIO

Figure 10: Tascam GigaStudio

The Nemesys GigaStudio (later **Tascam GigaStudio**) appeared on the market in 1997 and was the first software sampler on the market capable of handling libraries over 1 GB in size. It was also the first sampler implementing *disk streaming technology* on a large scale, which allowed to play back samples directly from the hard disk (loading just the beginning of each sampler in RAM) with extremely low latency. Already at the time, latencies lower than 3 ms were possible using high quality PCI-based sound interfaces (such as the RME Hammerfall).

The GigaStudio was also the first sampler instrument vastly outperforming any other sampler in terms of polyphony: **GigaStudio 160**, introduced in 2001, featured 160 voices of polyphony (allowing to emulate a full orchestra or a Grand Piano with activated sustain pedal without cutting any voice decay or release), which is about 5-10 times more the most powerful samplers available at the time.

With the sample library no longer limited to the amount of available RAM, it was for the first time possible to emulate large instrumental ensembles such as orchestras with extreme authenticity, as every note could be sampled separately at different dynamic levels, and for the first time at full length – without using loops.

CONTEMPORARY SOFTWARE SAMPLERS: KONTAKT, HALION, FALCON, MACH FIVE

The transition from hardware to software-based samplers began at the end of the '90s with the introduction of the Nemesys Gigastudio. Today hardware samplers have almost completely disappeared from the scene. Almost all modern samplers are software based.

Among the software sampler players / plugins available today are:

- **Native Instruments Kontakt** (Stand-alone, VST, AU): featuring probably the largest sample library available. today on any platform (including a lot of third-party developed libraries, like Heavyocity, Spitfire Audio, etc.)
- **Steinberg HALion** (VST, AU, AAX): including Synth (analogue-style synthesis), Sample, Grain (granular synthesis), Organ, and Wavetabel Synthesis type osicllators.
- **UVI Falcon** (AU, AAX, VST, Stand-alone): one of the most advanced samplers available today, developed in collaboration with IRCAM. Features 10 synthesizer-based oscillators and 7 sample-based oscillators (including classical, granular, slicing, stretch, etc.). 80 built-in effect modules, customable interface, scripting.
- **MOTU Mach Five** (VST, AU, RTAS, Stand-alone): a very powerful sampler, also based on the UVI engine

Most of these software samplers are available for the main audio plugin standards (VST, Audio Units, RTAS, AAX); they offer multi-mode filters, advanced envelopes and modulation options; can implement multisampling, velocity switching and crossfading; feature internal configurable effects, alternative oscillator types, and compatibility with the largest sampling libraries (AKAI, GigaStudio, EXS24, etc.).

Also worth mentioning:

- **Arturia CMI V**: a faithful software emulation of the Fairlight CMI
- **Arturia Synclavier V**: a faithful emulation of the NED Synclavier
- **Emagic EXS24** (only available in Apple Logic Pro): originally released in 2001, the interface is not quite state-of-the-art anymore, but it is still a very powerful sampler
- **112dB Morgana**: a sampler with an intuitive interface, designed to emulate the lo-fi sound of early samplers
- **Image-Line DirectWave** (VST only): a very complete sampler, from the makers of FL Studio.

Figure 11: UVI Falcon

Figure 12; MOTU Mach Five

Figure 13: NI Kontakt

SAMPLING TECHNOLOGY

The audio sampling technology is based on the principle of analyzing an analogue electrical signal a discrete (finite) number of times per seconds, with a device called ADC ("Analog to Digital Converter") and to store the value of the voltage as a digit (number) inside some form of memory device (RAM, disk, hard-disk, digital tape). During playback the values are read back from memory at the same clock speed and converted back into an analogue signal by the DAC ("Digital to Analog Converter").

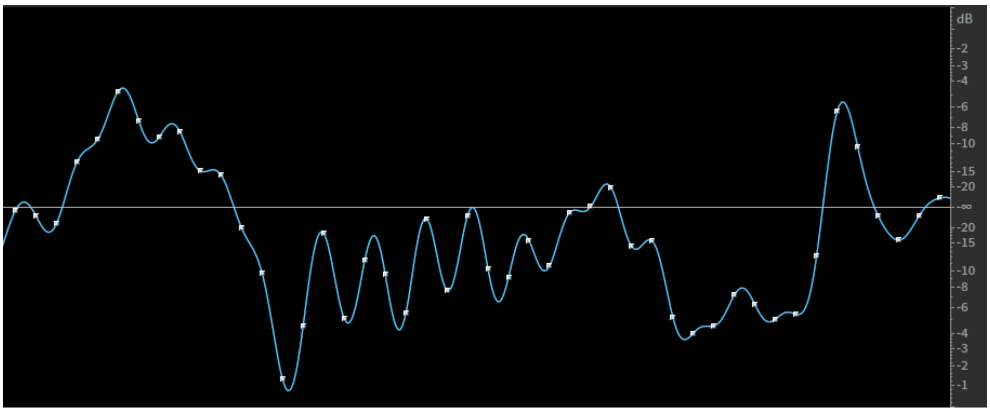

Figure 14: Audio Waveform display in the digital domain.
The samples are visible as white dots, while the band-limited analogue-signal
(as it would look, after DA conversion) is represented by the blue line.

Therefore, the accuracy of sampling is principally determined by these two parameters:

- *Sampling Rate* = how many samples are taken per unit of time
  (for example: 44, 1 kHz for CD-quality = 44 100 samples per second)

- *Quantizing Depth* = how accurate is every measurement
  (for example: 16-bit for CD-quality, allowing a range between 0 and 65 535 to be stored in memory)

Regarding the **sampling rate** (or sampling frequency), it must be noted that this must be *at least twice as much as the highest frequency to be reproduced* (Nyquist Theorem). Therefore, to achieve 20 kHz bandwidth (which corresponds to the range of frequencies we can hear) a sampling rate of at least 40 kHz is required.

In practice, to avoid a distortion called **aliasing** it is necessary to apply a steep low-pass filter to the signal just before the *Nyquist Frequency* (= half the sampling rate) before entering the ADC, and again at the DAC stage.

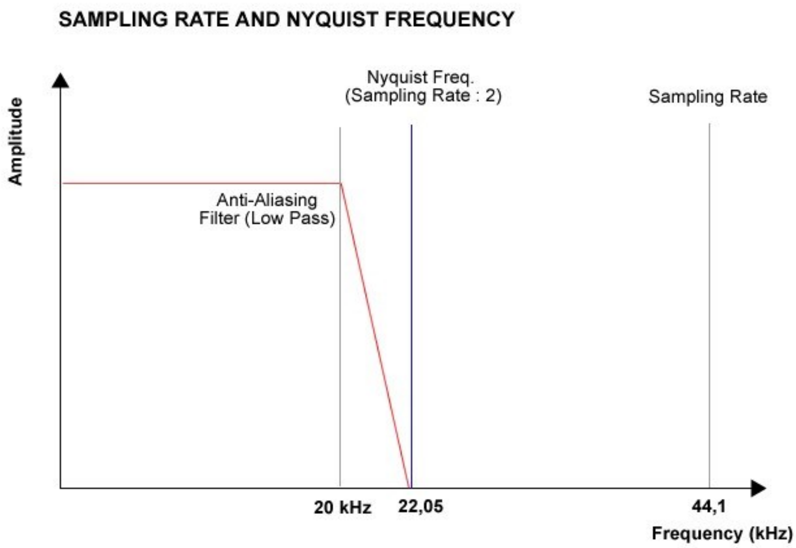

Figure 15: Sampling Rate and Nyquist Frequency

Without an anti-aliasing filter, signals higher than the Nyquist Frequency are incorrectly encoded by the ADC (basically they are “mirrored” around the Nyquist frequency), and appear again in the audible range as undesired signal components that have no harmonic relation with the original signal. This type of distortion sounds much more “artificial” than standard harmonic distortion, caused by analogue equipment.

For more information see http://www.digital-recordings.com/publ/pubneq.html

As low-pass filters cannot be infinitely steep, it is necessary to raise the sampling rate to more than double the desired bandwidth to ensure that no significant audio components are filtered out, and yet that no undesired components beyond the Nyquist Frequency pass through the ADC, as they would cause aliasing. Therefore, the typical standards for CD-Recording (44.1 kHz) and digital video or film (48 kHz) are both higher than 40 kHz.

44.1 kHz was the result of using video recorders as early digital audio devices – 3 samples were stored for each video line. 48 kHz corresponds to 2000 samples per movie frame.

To overcome the problem of using very steep analogue filters (which tend to create severe phase distortion in the signal close to the cutoff frequency), a technology called **oversampling** has been introduced: at the ADC stage the signal is only processed with a mild analogue low-pass filter that introduces very little phase distortion, and then sampled at a very high frequency (up to 128 times oversampling delta/sigma processing is used at times, corresponding to a sampling frequency of 5 644 800 Hz = 5,644 MHz).

Before downsampling the data stream to the desired target frequency (44.1/48/88.2/96/… kHz) for storage or further processing, a digital anti-aliasing filter is applied. In modern integrated circuit technology, it is much easier to implement such a steep filter in the digital domain, than as an analogue component.

A similar process can be applied at the DAC stage.

Quantization: just as the sampling frequency affects the maximum bandwidth of the recording, the quantization affects its dynamic range and the S/N (signal to noise) ratio. Each additional bit used to encode the signal adds about 6 dB dynamic range, so a 16-bit recording has approximately 96 dB, while a 24-bit has a theoretical maximum of 144 dB dynamic range. Other factors, such as the quality of the analogue equipment used for the ADC and DAC units can lower these values to a more typical 110-120 dB for 24-bit recordings.

High Resolution Digital Recordings use increased sampling frequencies such as 88.2 kHz (44.1 x 2), 96 kHz (48 x 2), 176,4 (44.1 x 4) and 192 kHz (48 x 4) and bit depths of 20, 24 or 32-bit float. The main advantage of higher sampling frequencies is to reduce the steepness of the anti-aliasing filter; this allows to increase the resolution and transparency of the higher frequencies (as there is little or no filter-induced phase-distortion).

Higher bit-depth translates into higher dynamic range and improved S/N ratio (lower quantization noise).

TYPICAL STRUCTURE OF A SAMPLER INSTRUMENT

A sampler instrument must be able to do much more than just digitally store a sound and play it back: it must provide real-time sample transposition over a wide frequency range, change the character of the sound with LFOs and filters, amplitude and filter envelopes, blend between different samples depending on dynamics, etc.

In some of its early implementations, samplers were very similar in structure to subtractive analogue synthesizers. For more information about a typical analogue synthesizer please check:
http://www.digitalnaturalsound.com/images/stories/fh_mma_courses/pdf/mg_sound_synthesis_01.pdf

In a sampler, the VCOs (Voltage Controlled Oscillators) are replaced by DWGs (Digital Waveform Generators), which read data from a digital memory (RAM).

The DWG must be able to transpose the sound to different pitches, as well as fine tune the samples. For transposition, either different playback speeds are used, or interpolation.

In early samplers, the output of the DWG was sent directly to a DAC, and from here on the sound path is completely analogue: the signal was then processed by an analogue filter and amplifier, either voltage (VCF, VCA) or digitally controlled (DCF, DCA). ADSR type envelopes modulated the pitch, the cutoff of the filter or the amplitude of the amplifier, just like in a typical subtractive synthesizer. These early samplers using real analogue filters and amplifiers had often a very pleasant and warm sound.

Later, with the development of better sounding digital filters and improved processing power of the DSPs (digital signal processors), the signal path was simplified as follows: DWG > Digital Filter > Digital Amplifier > DAC. While this was a much cheaper solution, it did not sound by far as “fat” and “warm” as the original analogue filters.

Also, the envelopes were simulated digitally, and were unfortunately very often not as fast (specifically, for percussive attack times) and accurate as the original analogue ones.

Only recently digital filter and envelopes have been developed to a point where they sound *as good as* or sometimes *better* than the original analogue ones. This also require considerable processing power, that was simply not available back in the ’80s and ’90s.

In early samplers, the same sample was often used across the complete keyboard range in order to use as little RAM as possible. This caused a loss of realism and fidelity, because the formants got transposed together with the note pitch. In most acoustic instruments there are certain formants that are *constant* across the complete instrumental range, because they depend on factors that do not change with the note being played. In an acoustic guitar, for example, certain formants depend on the body shape and size (which is the same for any note played), while others depend on the string thickness and length (that change with every note). In other words, many formants do *not* get transposed with pitch, which also means that each note has a slightly different tonal spectrum.

Multisampling overcomes this problem implementing different samples (each recorded separately at the original pitch) across the keyboard. Ideally – on tonally complex instruments, like a grand piano or saxophone – you would want to use a different sample for each key. In modern software samplers, that have virtually unlimited amount of memory thanks to disk streaming, multisampling can be implemented exactly in this way.

However, if the amount of RAM is limited (like in most hardware samplers) it might be enough to use about 4 samples per octave: each sample covers 3 notes and is never transposed more than 1 half-tone up, and 1 half-tone down, to keep the formant transposition from being noticeable.

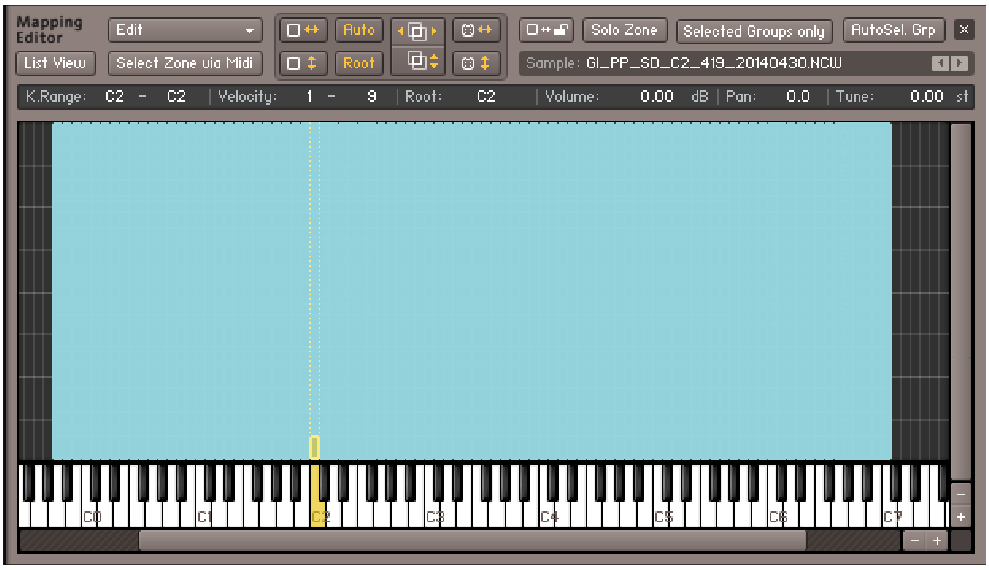

Figure 16: NI Kontakt: Keyboard Mapping Editor

In most natural instruments, a change in performance dynamics causes not only a change in amplitude, but also in sound color. For example, performing in a *piano* dynamic range results in a quieter sound with less overtones, while playing gradually more *forte* increases not only the level of the sound, but also the energy level of the overtones.

When using only one sample per note, these changes in color can partly be simulated by a low-pass filter (with just 6 or 12 dB /octave) that gradually opens the cutoff as the performance dynamics increase. For instance, you could sample a grand piano note performed in *fortissimo* and simulate all the other dynamic ranges (*forte, mezzo-forte, mezzo-piano, piano, pianissimo*) by modulating both the amplitude level, as well as the filter cutoff and/or envelope amount in order to create gradually quieter as well as duller sounds (less rich in overtones).

Velocity-Switching or *Velocity-Crossfade* allow a more authentic reproduction of the sampled instrument dynamic range: this is realized implementing for each note multiple samples, recorded at different dynamic levels or with different playing techniques. The playback engine either switches between samples (each sample covering a certain dynamic range), or crossfades gradually from one to the next, based on input keyboard velocity.

Looping (single cycle, range, crossfade, back & forward) was introduced already in the early sampler instruments in order to spare the amount of required RAM. Only the initial portion of a note would be sampled in full, then a section (either a single cycle wave, or a longer portion of the wave using crossfade to avoid clicks) would be looped and – if the instrument was a decay-type, like piano or guitar – gradually decreased in level using an envelope, to simulate the natural decay of the instrument.

Some software samplers such as Native Instrument Kontakt, Steinberg Halion or UVI Falcon also offer alternative types of oscillators (analogue-like, granular, wavetable, etc.), an integrated digital mixer with insert and send/return buses, and integrated effects (reverb, delay, saturation, chorus, etc.).

LINKS

- Mellotron http://www.mellotron.com
- Hammond Organ http://theatreorgans.com/grounds/docs/history.html
- Vintage Synth Explorer: www.vintagesynth.com
- Fairlight CMI http://www.vintagesynth.com/misc/fairlight_cmi.php
- Fairlight CMI (review) http://www.soundonsound.com/sos/apr99/articles/fairlight.htm
- NED Synclavier http://www.vintagesynth.com/misc/synclav.php
- EMU Emulator http://www.vintagesynth.com/emu/emulator.php
- EMU Emulator II http://www.vintagesynth.com/emu/emulator2.php
- EMU Emulator III http://www.vintagesynth.com/emu/emulator3.php
- Ensoniq Mirage http://www.vintagesynth.com/ensoniq/ens_mirage.php
- Ensoniq ASR-10 http://www.vintagesynth.com/ensoniq/asr10.php
- Roland S-50 http://www.vintagesynth.com/roland/s50.php
- Roland S-770 http://www.vintagesynth.com/roland/s770.php
- Korg DSS-1 http://www.vintagesynth.com/korg/dss1.php
- AKAI S900 http://www.vintagesynth.com/akai/s900.php
- AKAI S1000 http://www.vintagesynth.com/akai/s1000.php
- NI Kontakt – Creating a new Instrument https://www.youtube.com/watch?v=om-aKoLS58U

WEBSITE

- http://www.digitalnaturalsound.com/fh--multimediaart/music-production.html
- www.digitalnaturalsound.com or www.dns-studios.com > FH | MultiMediaArt > Music Production