# Cutting Tracks, Making CDs: A Comparative Study Of Audio Time-Correction Techniques In The Desktop Age.

Justin Paterson

London College of Music, Thames Valley University
justin.paterson@tvu.ac.uk

## Abstract

*Producers have long sought to rhythmically 'tighten' studio performances. Software-based DAWs now come with proprietary functions to facilitate this, but only the latest generation of platforms allow relative ease of use on longer takes. Each method has advantages and disadvantages in terms of ease and speed of use, transient preservation, implied subsequent workflow and (usually) unwanted artifacts. Whilst rhythmically consistent material with clear transients is readily controllable with contemporary tools, working with complex mixtures of note-values still presents a challenge and requires much user intervention.*
*This paper performs a comparative study of different audio quantize techniques on percussive material, often on rhythmically complex performances. It will seek to compare necessary methodologies and workflow implications through the use of several contemporary systems: Recycle, Pro Tools, Logic, Cubase, Live, Melodyne, and Nuendo. The current level of man-machine interaction will be explored, and the best features from each platform will be collated. A model for the future will be speculatively presented.*

## 1 Introduction

The application of various techniques to adjust the musical timing of studio performances is commonplace today, and much debate rages over the merits of various commercial systems, particularly concerning their workflow, and effectiveness. Most professional practitioners are familiar with one, two or perhaps three Digital Audio Workstations (DAWs) and possibly a number of other legacy systems (which do not necessarily represent that system's current evolution), and subsequently this debate tends to be skewed towards the experience of the individual, perhaps even with a partisan approach. This paper will compare seven current systems that offer time correction features, thus offering a broader perspective on this debate. Of course, various other systems also exist, each with their own merits and limitations.

Workflow is a crucial part of any time-correction system. The days of tape editing gave way to a numerical digital approach, and this in turn was superseded by graphical methods, which still represent the state-of-the-art. Consideration of this chronology immediately reveals the relevance of the workflow issue since it is most apparent that these approaches yielded different quality of results with different amounts and types of user mediation. The paper will analyze and demonstrate crucial and idiosyncratic procedures of contemporary desktop systems through multimedia content. This will enable the reader to assess how systems with which they might not be familiar address typical time correction procedures and hence make a judgment on their efficiency. It will not simply form a list of subjectively comparative sonic qualities such as that demonstrated by Guitton (2008), which would entail a separate study, and so 'quality' will be largely ignored except for when the process introduces severe artifacts, although the multimedia content will allow readers to form their own judgments. It should be emphasized that the procedures demonstrated here are not necessarily executed to their maximum effect- it is the process which is relevant. These processes may have previously been described in numerous ways, usually by consumer periodical reviews, but it is the context of comparison that forms the bulk of this work. Advantages and disadvantages of each system's workflow will be highlighted. Without attempting to produce a feature-set, but rather by highlighting operational issues that contrast/compare with the other systems considered, the 'Workflow Implications' offer a summary relevant to user experience.

In addition, those features deemed best from each system will be collated into an 'idealized' specification for a future system, and this system will be discussed. The intention here is not to design a functional system, but simply focus the reader on how tomorrow's producer might aspire to work.

### 1.1 Current Approaches

The term 'audio time-correction' can be taken in this context to mean the adjustment of the timing of individual notes of the audio recording of a performance towards an idealized placement. The definition of 'idealized' is beyond the scope of this paper, and issues of groove, in /out of time and their associated ethics will not be discussed. There

currently exist two basic approaches to audio quantization: Slicing and Stretching.

1] Slicing typically involves defining an area of a 'take' (a contiguous area of disk space containing a recording) to be its own 'region' (a reference to a smaller section of the take), and this new region can placed on the time-line of the DAW to align its preferred note transient to user preference. Such an approach can be applied to as many notes in a performance as required, and the computer simply has to randomly access portions of the original region according to the new timings implied by the slicing.

2] Stretching allows movement of a given note's transient on the time line, and subsequently applies a time-stretch (compression or expansion) algorithm to the areas before and after to compensate for the movement. This stretch might be applied in real time, or the audio might be rendered into a new region.

The systems considered in this paper either offer one or both of these procedures.

In addition, some of these systems are optimized for operation on 'loops' (portions of audio typically only a few bars in length), and some are capable of operation on takes of any length.

## 1.2 Test Material

As stated, this paper will focus on purely percussive material; this is to simplify the appraisal since pitched material presents many more scenarios. When slicing, percussive material tends to offer the perception of a 'gap' between the relatively short transients with the exception of naturally sustaining sounds such as cymbals, and distant mic'ed recordings such as room mics on a drum kit that typically contain continuous ambience between notes. Pitched material will more often sustain between notes, and so the slicing will tend to create more artifacts. Stretching is also more complicated since typically there will be numerous algorithms available for operation on different types of monophonic and polyphonic material, and so the number of permutations to be described goes beyond the scope of a single paper.

Two principle audio sources are operated on: a two bar drum loop in 4/4 with essentially a 1/16 note feel (this is utilized both with and without room mics), and a hi-hat part[1] containing many different note values in a swung feel, four bars of 7/4 in length.

Play hi-hat

Play Drum Loop

[1] This hi-hat part was crudely recorded for a demo. A long-term physical injury prevents it being played more precisely for the final production, and an attempt to salvage it provided the inspiration for this work.

# 2 Case Studies

There now follows a number of examples of typical processes, capabilities and workflows in the chosen systems. Video examples demonstrate the user experience and idiosyncratic advantages and disadvantages are highlighted. Inevitably, system-specific terminology is applied as necessary.

## 2.1 Recycle

**Overview:** Recycle was released by Propellerheads in 1994 (Propellerheads, 2004). It signified a change in the way musicians worked since for the first time in common usage, it allowed a monitor-sized (as opposed to small hardware sampler displays) graphical visualization of the audio and had auto slice detection/stretch based on transient analysis. It also featured various features to manipulate tempo and pitch independently and envelope shaping of individual slices. The workflow it presented has loosely remained industry standard to the present, and has been emulated to varying degrees by various other manufacturers.

**Advantages:**

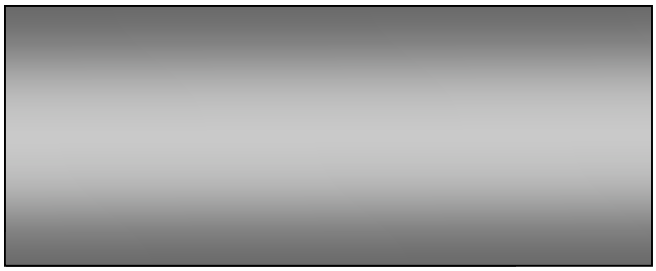
- Ability to lock multiple slices in edit mode
- Good auditioning for individual slices
- Adjustable stretch factor, independent of transients
- Multiple output format, e.g. (native) .rx2, .wav+.mid
- Optimized (variable) tempo output
- Good transient detection
- Individual slice output to a sampler offers very good groove flexibility, and comprehensive control of slice order
- Direct DAW import of .rx2s may allow the setting of an auto cross fade or various track options, e.g. with Logic: render into a single loop, Apple Loop or split over numerous DAW tracks per slice.
- Running a complex set of note values from a sampler allows the generally more detailed MIDI editing features of a DAW to be employed, e.g. easy changes of grid.

**Disadvantages:**

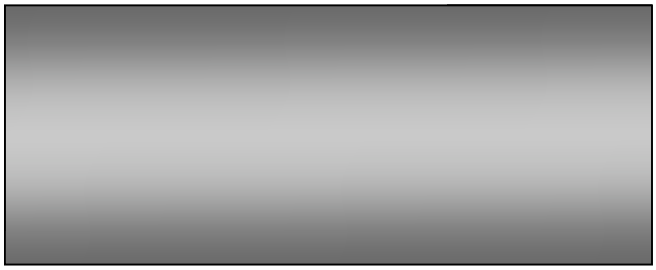
- Only 92 slices available which limits its use to 'loops' rather than audio takes
- No relative time adjustment of notes available
- A more flexible multi-function tool would be faster, e.g. sensitivity adjustment via modifier-drag instead of the fader, and pencil with audition and lock all from a single tool
- Not integrated into a DAW

- Individual slice output to a sampler (which is required) means labouring with identical looking MIDI notes at different pitches to edit a sequence
- Often requires a very iterative workflow to get good results
- Direct import of .rx2s into Logic as audio can result in strange behaviour with the need to adjust anchors, remove overlaps, and still results in glitches when say a crash crosses a ghosted snare transient which carried a slice point

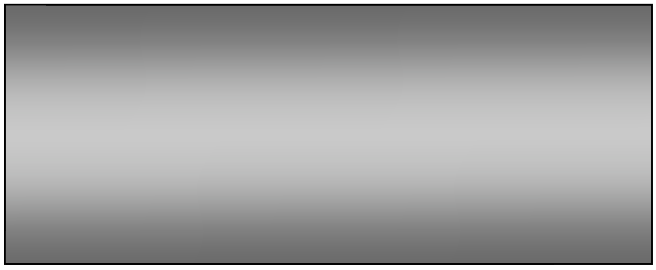
**Examples:**



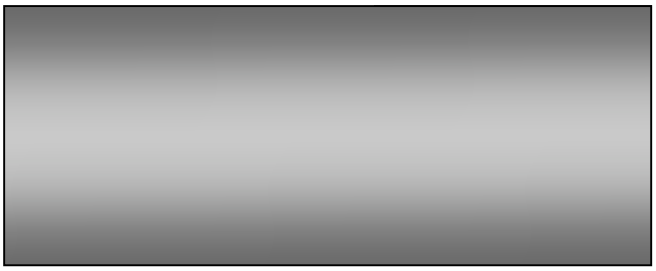
Movie 1: A demonstration of the iterative detection of slices in Recycle.



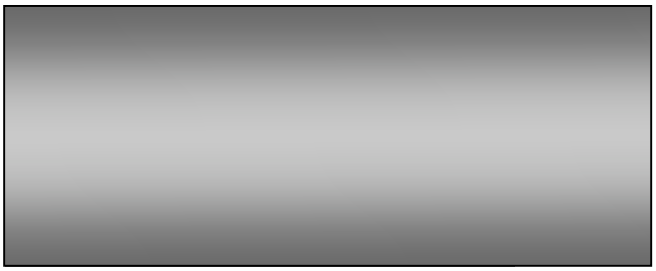
Movie 2: The relationship between the stretch factor and the tempo using Recycle's preview mode.



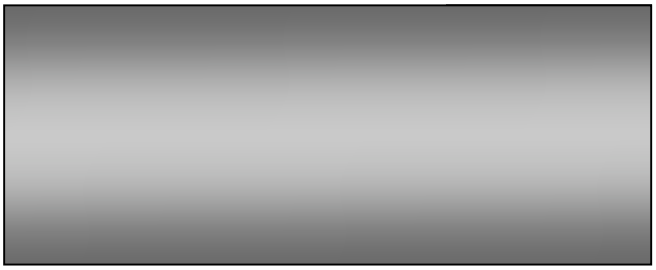
Movie 3: Importing an .rx2 into Logic and compensating for a tempo drop.



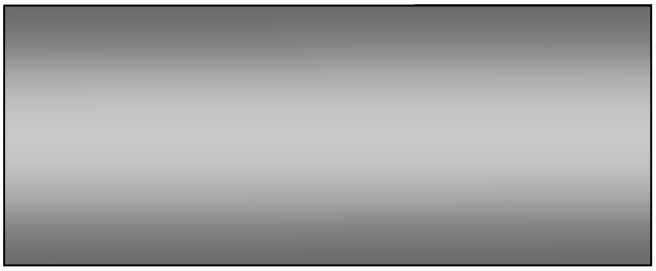
Movie 4: An illustration of slicing problems common to all such systems, but here demonstrated through the import shown in Movie 3.



Movie 5: Recycle's MIDI file has been imported into Logic, and the quantization in the MIDI domain is demonstrated.



Movie 6: Applying a shuffle to Movie 5 groove. Note that the anomaly mentioned necessitates going back to Recycle and re-editing the offending slice.



Movie 7: The Hi-Hat has been imported and quantized in the MIDI domain.

**Workflow Implications:**

- The user will tend to use loops; extra effort is required to cut whole audio takes into smaller sections
- Much material requires an iterative approach to defining slices

- The multi-function tools could be more comprehensive; changing tool is still required for certain operations
- It is necessary to leave the native DAW environment
- There are very creative possibilities through 'rephrasing' the loop through MIDI editing in the DAW since such editing is often more flexible than the equivalent in the 'pure audio' domain

Recycle is a formidable system that has defined most subsequent systems' workflow. Post-slicing, working in the DAW MIDI domain does not suit all situations, and the MIDI note-driven finite loop length is limiting. Direct import of .rx2s can yield good results within certain constraints.

## 2.2 Cubase and Nuendo

**Overview:** Steinberg's Cubase VST offered audio stretching (rendered) to hit points since its inception in 1996, and SX2 offered 'Timewarp' in 2002 allowing a radical stretching of audio within the DAW environment. Nuendo is considered alongside Cubase here, since the front end GUI is for the most part identical, and both will be referred to as 'Cubase' forthwith. Cubase is a high-end DAW with an extended feature set for both MIDI and audio applications. Current incarnations have comprehensive slicing and stretching capabilities; Bachmann et al. (2006) and Knowledgebase.Steinberg (2008) state that both now use the MPEX3 algorithm for stretching, although for monophonic material, Z-Plane.development (2008) state that élastique SOLOIST is used in Cubase 4. Steinberg refer to stretching as Warping.

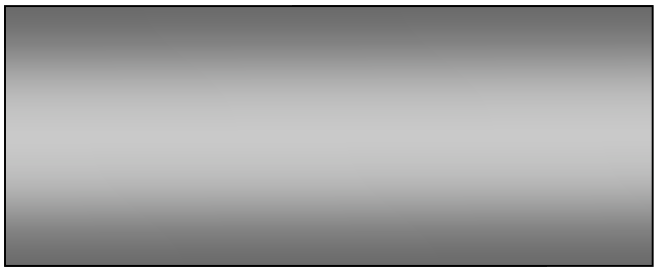
**Advantages:**

- 'Metric Bias' gives Hitpoints close to even meter divisions a sensitivity boost, which raises their priority relative to the sensitivity slider
- Slices can be transposed (and enveloped) individually, which can be used very creatively: by semi-tones for harmonic changes, or in cents for tonal corrections or subtle percussion variations
- The 'Quantize Setup' dialogue works nicely for simple variations
- Once audio is sliced into component Events within Parts, manipulations including 'order of events' (phrasing), Snap-points, fade-ins/outs, volumes and quantize all become easily controlled, and the Part acts as a 'folder' allowing easy musical arrangement
- Manually setting Hitpoints is tactile and similar to Recycle. Easy to hear the end of auditioned slices to gauge if the next slice is well timed
- Audio waveforms 'scale' horizontally as Warp Tabs are adjusted
- Using Warp Tabs created from Hitpoints is effective and tactile, with even apparently over identified tabs working well when just applying Audio Quantize from the menu
- When snap is on, the magnetic feature allows quick adjustments of 'obvious notes'
- A degree of multi-functionality for the editing tool
- Auto cross-fading of regions
- A strip silence then snap to events allows manual cutting of a multi-track folder

**Disadvantages:**

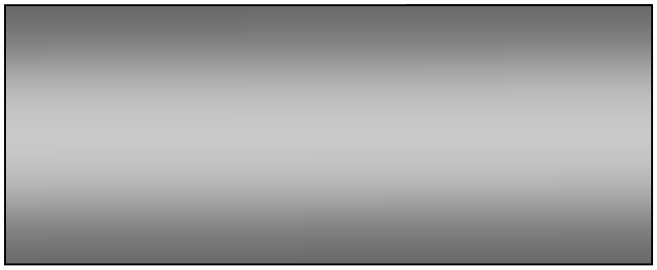
- Locking Hitpoints is tedious, since they generally have to be done individually with little visual feedback
- 'Snap to Zero' allows click-free slicing, but can compromise absolute accuracy
- 'Slice and Close' tends to create noisy gaps when dropping the tempo, and unfriendly grey areas of overlap when raising it
- Warp Tabs work in real-time, but require that the part is Flattened (rendered) to optimize sound quality with MPEX3
- 'Quantize Setup' supports tuplets, but the un-expandable graphics do not make this clear as the display of capture zones overlap
- Stretching is not always as accurate as one might hope. The 'Flatten' function in particular could not be made to preserve hi-hat transients adequately; performing a mixdown is better
- Manually setting Hitpoints is time consuming
- No key command to sequentially audition slices; they all have to be clicked
- A second pass at a higher zoom is often required to tidy errors
- No adaptable grid available when viewing hitpoints, which makes many operations impossible
- When snap is on, the magnetic feature does not allow very small edits.
- It can be quite arduous to change quantize value constantly for rhythmically complex material

**Examples:**

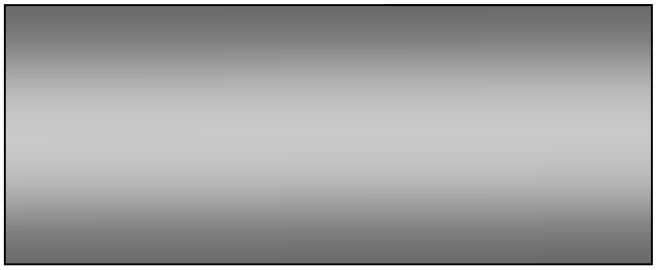


Movie 8: The process of locking Hitpoints on part of the Hi-Hat loop. These have to be done individually before the sensitivity can be reduced.

Movie 9: The process of quantizing the user defined Hitpoints in Cubase on part of the Hi-Hat loop. Note that as quantize is applied, Cubase disregards some that were already defined, although with little detriment.



Movie 10: Manual adjustment of Warp Tabs on a different section of the loop. Both compensation for slightly erroneous detection and magnetic snapping can be seen.



Movie 11: The drum loop has been imported into Cubase, which has a song tempo 10bpm below the loop's 110bpm. Quantizing using Warp Tab markers requires some user mediation. Note again the magnetic snap, but also how the grid is not visible as an aid.

Movie 12: Here, the room mics of the drum loop are included. The Tempo is reduced, and automated slicing is applied. The ambience is disrupted, despite an attempt with cross-fading.

Movie 13: The sliced dry loop at 100bpm has a progressive swing applied. Note how a value of swing is set, and then applied retrospectively. This inhibits ease of gradated auditioning.

**Workflow Implications:**

- When slicing, considerable mediation is required to manually adjust fades and overlaps of regions to attempt smooth playback
- Much material requires an iterative approach to defining slices
- When warping, it is sometimes necessary to 'estimate' target positions of transients due to the lack of grid
- The warping process is extremely tolerant of both extra warp tabs and extrapolating unspecified transients between encompassing tabs, which can lead to slightly lateral phrasing
- Interesting creative possibilities through the individual slice envelopes and their transposition options
- Suitable for both loops and takes
- Combinations of note values require toggling of quantize settings

Cubase and Nuendo offer a sophisticated feature set, which with sufficient mediation can yield good results on complex material. Either slicing or stretching can be utilized as deemed appropriate. Straightforward material of a nature aligned with key features is relatively intuitive/easy to deal with.

## 2.3 Logic Pro

**Overview:** In 1997, Emagic introduced Logic Audio 3, the first version to allow non-linear editing of audio. The current Apple version is a full-featured DAW, comparable to Cubase and Nuendo. Slicing is possible, and stretching uses native Apple algorithms, although third party algorithms from Izotope and Serato can be integrated seamlessly. Logic features a native 'ACID-like' format called Apple Loops, but although the Apple Loops Utility application allows easy user-detection of transients a la Recycle, there is no facility to adjust relative timings of notes either within itself or once imported into Logic.

**Advantages:**
- Strip Silence automates slicing for whole Regions (takes).
- The Marquee tool can auto-detect transients (one at a time), which aids manual slicing
- Discrete regions can be comprehensively quantized
- Auto cross-fading of regions
- The Quantize Engine features a similar interface to Beat Mapping
- Accepts 3rd Party Time Stretch algorithms
- Manual Trim (stretch) tool can compensate for user-created gaps between regions
- Can create MIDI renditions of audio

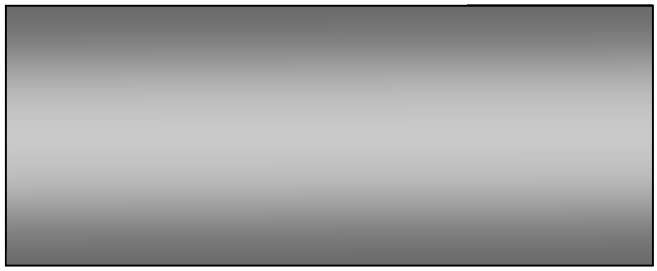
**Disadvantages:**
Strip Silence:
- Not suitable for highly ambient sounds– it requires a degree of silence (or at least significant dynamic range) to function
- The window is not zoomable which forces the user to iteratively operate on sub-sections of longer files, which becomes a highly protracted procedure unless very dry material is being processed
- The use of numerical parameters are not very tactile

Marquee:
- Does not detect all transients
- Awkward default key commands

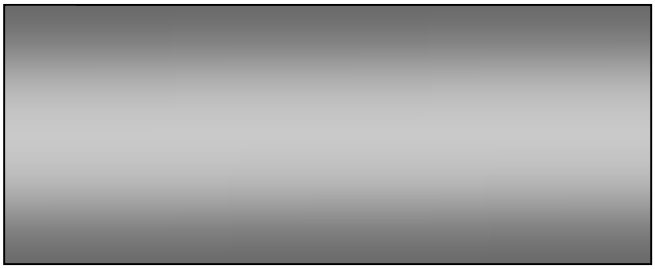
Quantize Engine:
- Unfriendly numerical parameters
- Virtually un-editable detection system– less or more, with no locking of detection points or transient separation
- Destructive (since it is a part of the Sample Editor)
- Easy to induce random digital crackles during operation

Both:
- Quantizing in the numerical Event List is awkward

**Examples:**



Movie 14: Slicing, using the Marquee tool to detect the transients in Logic. The slice artifacts can be heard, and with subsequent editing can be improved, but this is typical of the initial result. Use of 'Strip Silence' produces similar issues.



Movie 15: Logic's 'Quantize Engine' easily detects the transients, but the native 'Drums Mid' algorithm struggles.

**Workflow Implications:**
- Work with loops or takes
- Much mediation is typically required for slicing
- 'Audio Quantize' is best used only on isolated short regions
- Strip Silence works best on short, dry staccato recordings
- Combinations of note values require toggling of quantize settings

Logic Pro is capable of a range of quantize approaches, but requires much mediation to achieve the best results. Some users have developed ingenious techniques for say, multi-track drum editing as can be seen in Pye (2008), but the convoluted approach is compensatory for first order functionality in the system.

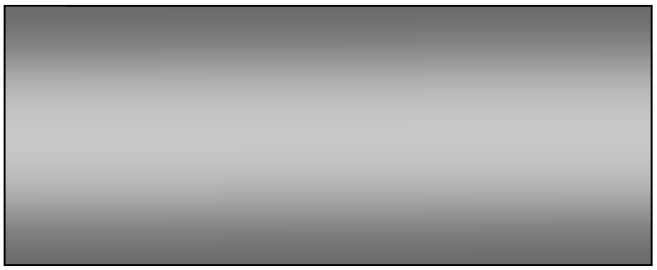
## 2.4 Live

**Overview:** Ableton Live was released in 2001, and from the onset was targeted towards audio quantizing. More often marketed as a DJ type tool, it also offers formidable all round DAW functionality and can both slice and stretch (Warp), although it is optimized for Warping, for which it employs Z-Plane's élastique efficient v1.3 algorithm as stated at Zplane.development (2008). It allows all quantize operations to be performed in real time, although material can later be rendered if desired.

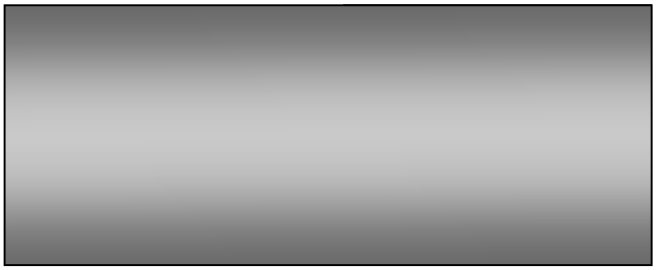
**Advantages:**
- Very tactile placement of markers
- Real time warping at optimum quality
- Good preservation of transients in Beat Mode
- An auto-adaptive grid

**Disadvantages:**
- Other than a crude global setting, Live does not do cross-fades, so slicing is not an easy option
- User defined transients are labour intensive; one at a time
- Warping does not scale the waveform dynamically, so it can be hard to read the time-stretched numerical grid; this can lead to placing warp markers in the 'wrong' place during intricate editing
- The tactile nature of the warp-editing makes it very easy to introduce unwanted artefacts. This produces a dichotomy between rhythmic accuracy and tonality
- Lack of a transient to cursor function means that slicing individual notes is arduous
- Setting swing 'on' can degrade certain audio considerably and is applied globally, so for example, loosely swung 16ths and actual triplet 16ths compete

**Examples:**



Movie 16: Live; using Warp markers to correct the Drum Loop.



Movie 17: The warping process to correct two sections of the Hi-Hat Loop is demonstrated.

**Workflow Implications:**
- The warping process is extremely tolerant of both extra warp tabs and extrapolating unspecified transients between encompassing tabs, which can lead to slightly lateral phrasing; this is likely due to the independence of the transient analysis from the visuals of the markers
- Despite sophisticated auto-detection of phrasing, much mediation is required to identify and correct relevant transients
- The best results for complex combinations of note values come from manually identifying transients and adjusting the grid accordingly, which is labour intensive
- Suitable for both loops and takes

Live offers an idiosyncratic and powerful approach to audio quantizing and can also deal with multi-track scenarios effectively by imposing warp adjustments onto a number of target tracks. Again, much mediation is required to quantize complex combinations of note values. Its stretch artifacts can be used creatively in modern 'Glitch' styles.

## 2.5 Pro Tools

**Overview:** Many have considered Digidesign Pro Tools as the industry standard non-linear recording platform, perhaps since the release of Pro Tools III in 1994. In 2003 the 'Beat Detective' auto-slicing tool was introduced for TDM systems. In 2006, iZotope licensed their Radius algorithm to Digidesign (Izotope 2006) for off-line processing branded as X-Form, and in 2007, 'Elastic Audio' was implemented in v7.4 using principally native algorithms in addition to X-Form. The Elastic Audio approach offers a choice of several algorithms, and a Recycle-style transient detection/editing.

**Advantages:**
Beat Detective
- The unified interface is the same for groove extraction/slicing/tempo extraction in both MIDI and Audio
- 1/2 speed playback is good for hearing timing anomalies (shift/space-bar)
- 'Scroll Next' is handy when working at high magnification
- Can work on multi-tracks via 'Collection Mode', progressively adding unique beat triggers from various regions and then separating with all relevant multi-tracks selected
- 'Edit Smoothing' works well
- 'Region Conform' includes the quantize parameters: Strength, Exclusion and Swing

Elastic Audio
- Linear time base with graphical scaling of waveforms
- Ability to control decay in the rhythmic algorithm, and various parameters in others
- Good (modifier driven) multi-function editing tools to control Warp Markers
- Excellent auto-detection of transients
- Works on multi-track audio too via groups
- Very smooth workflow

**Disadvantages:**
Beat Detective
- Cannot see Promoted/Demoted slices without changing the sensitivity
- Cannot 'one-click' audition individual slices; scrub is necessary
- Cannot see the grid superimposed on the audio

- Sometimes necessary to iterate settings to achieve successful results
- The usual overlap-related slicing problems: transients within sustained sounds are messy. 'Edit Smoothing' can help, but is still hit and miss, and is especially poor at applying a groove which necessitates a larger time shift, although further interaction with the time stretch tool can help

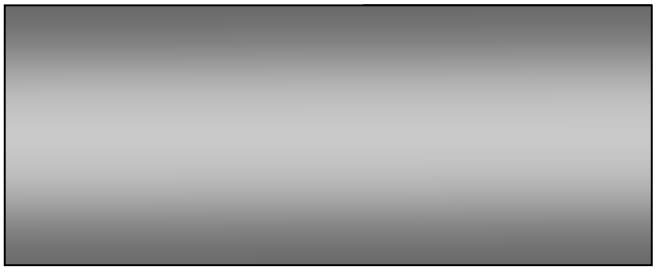
Elastic Audio
- Decay works well for dry drums, but not so well with reverb
- It is clumsy to work with an embedded groove feel (e.g. 16D) and also sometimes conform to grid values e.g. 16 triplet– it is necessary to toggle between modes by mouse/menu
- Cannot see the grid superimposed on the audio
- It takes 2 visits to a menu to toggle between e.g. triplet 1/16s and 1/8s
- X-Form algorithm is only available as a rendered process

**Examples:**

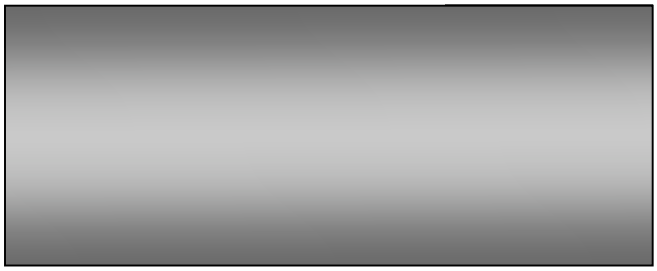
Movie 18: The Pro Tools 'Beat Detective' workflow is demonstrated. All stages of the edit take place in the native Beat Detective window.

Movie 19: Beat Detective now applies a (Logic style 16D) shuffle to the Drum Loop. The slices reach the limit of their effectiveness without further user-intervention; typically stretching of individual slices, which is not demonstrated here.

Movie 20: Pro Tools 'Elastic Audio' is applied to the same Loop to quantize first straight, and then 16D.



Movie 21: The 'Decay Rate' parameter of the Rhythmic algorithm is applied to the Hi-Hat Loop. This is akin to what has to be done natively within Recycle, although clearly, the ability to do this within the DAW is an advantage.



Movie 22: The Hi-Hat Loop after quantizing with the various note values in 'Elastic Audio'.

**Workflow Implications:**

Beat Detective:
- Mediation may be required to manually adjust fades and overlaps of regions, to attempt smooth playback
- Dry recordings work better
- Combinations of note values require toggling of quantize settings
- Problems with sounds which sustain through other transients

Elastic Audio:
- Combinations of note values require toggling of quantize settings
- The optimum algorithm must be selected, along with manual adjustment of the parameters in rhythmic/polyphonic plug-in windows
- Very transparent workflow

Beat Detective offers excellent automation of the slicing process with a workflow contained within its own window, and produces good results on certain kinds of material, but requires intensive mediation on others. Elastic audio has excellent transient detection, and makes the quantize process very easy on rhythmically consistent material, although mixed note values still require much toggling of values.

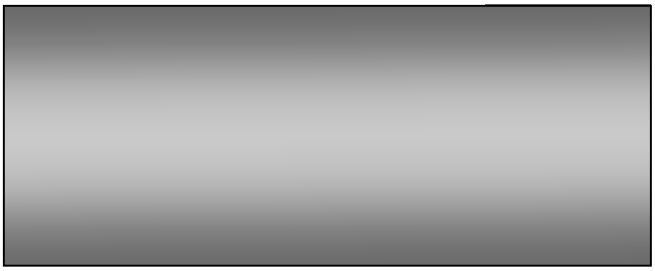
## 2.6 Melodyne

**Overview:** Celemony's Melodyne was introduced in 2001, and was marketed principally as a pitch adjustment/transposition tool. It uses a native algorithm called 'Local Sound Synthesis' (Hoenig and Neubäcker 2005, p.24), and is quite capable of quantization operations. At the time of writing, the new DNA system that can separate voices of a chord and treat each individually is eagerly awaited, but it remains to be seen if this can deal with voices of different timbre, e.g. stereo drum loops.

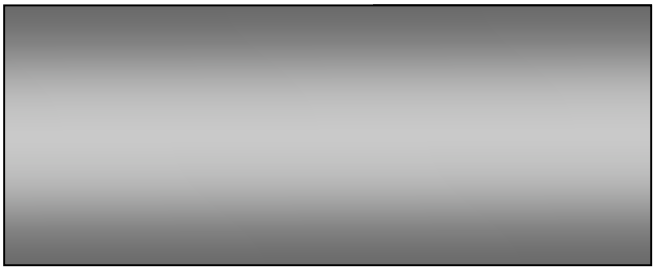
**Advantages:**
- A tactile editing system with a set of mouse tools that allows very fine adjustments
- Good quality time stretching which allows very naturalistic phrasing
- The Quantize dialogue moves blobs and grid in real time, allowing good visual feedback
- There is an intelligent quantize algorithm that allows larger note values to be set, and phrases shorter note values within that framework
- Excellent detection of transients using pitched material without a defined amplitude attack
- Easy to adjust other parameters e.g. volume of individual notes
- Can create MIDI renditions of audio

**Disadvantages:**
- Quantization options are fairly crude in terms of pure note value options, but it does offer a quantize amount parameter, 0–100%
- Transients are hard to discern visually due to the rounded graphic style of the 'Blobs'
- Not very rigorous auto detection of transients, although much more sensitive to pitched material
- Large amounts of user interaction required; time consuming
- Cannot audition a single slice to check for visually hidden transients

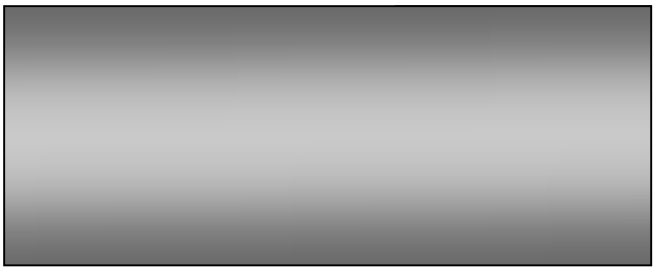
**Examples:**



Movie 23: Activating 'Quantize' on the Drum Loop in Melodyne has an instant effect, and then visual feedback is demonstrated during progressive increasing of quantize strength.



Movie 24: Detailed slicing is demonstrated, and the visual/aural guessing associated with applying a non-triplet shuffle is shown.



Movie 25: The quantized Hi-Hat loop is played. Beat 5 of bar 4 could still benefit from the attention of Melodyne's volume tool.

**Workflow Implications:**
- Principle transients can be auto-detected, but mediation may be required to identify others
- The tools offer much flexibility and a degree of context sensitivity
- Although grooves can be transferred between tracks in Melodyne Studio, it is awkward to develop in a single track– the grid/snap is fixed
- Tactile adjustment of quantize strength
- Much mediation is required for complex passages, although more basic ones are straightforward

Melodyne is an extremely powerful and creative quantization tool, although there are still areas that require extensive mediation. Its other (pitch/formant/volume) features can greatly enhance the final result too, and abuse of its feature set can produce remarkable results.

# 3 The Future: A Hybrid Model

After the appraisals of Section 2, the best features from each system can be extracted, combined and extended to indicate a specification of an idealized future workflow design. The design is not intended to be definitive, but merely suggestive. The specification will be partitioned into three areas: Interface, Operation and Feature Set.

## 3.1 Interface

**Multi-touch interface (with voice control?).** Any speculative design must look beyond the mouse and keyboard as Han (2006) has demonstrated so iconoclastically.

**Colour coding.** Traditionally, MIDI pitches/velocities etc. utilized this, but too little use is made of colour in audio applications.

**Auto-adaptive grids display grooves visually.** Live's adaptive grid concept can be extended to include groove as well as 'traditional' note values, and both should be available in a regionalized form in hybrid combinations, superimposed upon the waveforms.

**A unified interface for groove extraction, slicing and tempo extraction in both MIDI and Audio.** Users do not wish to learn too much, so a sharing of GUI/workflow for multiple processes is essential. Pro Tools Beat Detective and Logic's Beat Mapping/Quantize Engine are current examples

**Integrated Stretching and Slicing with a shared GUI.** Current systems present stretching and slicing as separate approaches. Slicing has many sonic advantages (currently), yet it is heavily limited by truncation of natural decay and transients within sustained portions. Stretching can appear more fluid, but is more prone to artifacts. Integrating these intelligently and invisibly would allow the user to always work with the optimal technique.

**Switchable linear time base/Linear Note value grid.** The Cubase/Pro Tools linear note value and the Live linear audio time-base should be switchable as the user requires.

## 3.2 Operation

**A comprehensive, context dependant set of 'mouse' tools, including slice sensitivity control.** One of the things which slows workflow is changing mouse tool. Although many current systems implement this partially, this should be implemented comprehensively via multi-touch type modifiers.

**A palette of user selected grids with linked quantize a la old Cubase, and audio reference click subdivisions, colour coded.** Typical operations do not need instant access to all possible note values and grooves, so a template of the most common for a given passage should be easily available. 'Old Cubase' refers to a discontinued function that allowed the grid and quantize value to change together. The reference click should reflect the subdivisions defined in the auto-adaptive grid to allow the user to hear idealized timings for a particular section to better choose a quantize value.

**Continuously controllable quantize (via hardware) strength/groove/strength and decay/grain etc for tactile control with visual feedback.** Quantize strength etc. should be controllable via assigned hardware/virtual controllers, and be available at all stages of the workflow. Facilitating 'MIDI' type control would allow this to be modulated/automated in real time

**Cubase/Nuendo-style 'Metric Bias' giving markers close to even meter divisions a sensitivity boost, along with a magnetic slice-to-marker option.** This should be easily toggle-able to allow fine editing when necessary.

**'Scroll Next' for working at high magnification, with zoom presets centred around the selected marker.** Pro Tools 'Scroll Next' and zoom buttons should be combined with a multi-touch zoom as demonstrated by Han (2006).

## 3.3 Feature Set

**Long 'phrase' groove templates with different note values.** Phrases should be of any length and feature 'cadence points' where the required note can be locked to the desired position as in Logic's Beat Mapping, with Melodyne intelligent (user-driven) analysis for notes in between. This would be good with bass and drums.

**A Logic Beat Mapping style tempo map to integrate tempo variations superimposed upon quantize.** This should share the GUI with the above 'phrase' feature.

**Melodyne DNA style separation of components within a 'slice' to render individually quantizable elements.** Problems frequently arise when flams on different kit components are embedded in tracks. Separation of these could solve this problem.

**Individual transposition and enveloping (beyond 'just' volume) of slices.** Cubase's Event envelopes might be developed to include Melodyne variable parameters and more, as implemented in devices such as Stylus RMX.

**Real time warping at optimum quality.** A number of current systems do not offer this.

**1/2 speed playback.** Pro Tools' simple but effective function can greatly aid the ear in resolving desired timings.

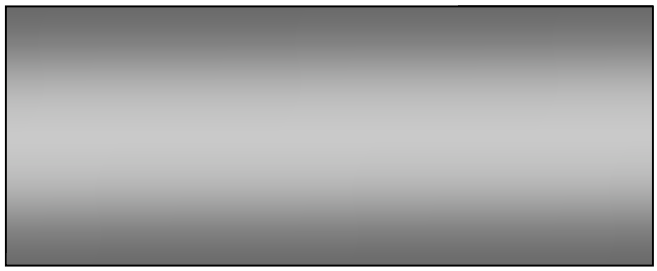
**Pro Tools style 'Collection Mode'.** Drawing data from multiple tracks to form overall templates for adjustment.

**The ability to create polyphonic MIDI renditions of audio.** Zenph Studios (2008) are already implementing this technology for piano music, and DNA promises to augment this. The resynthesis of source material under such control would provide many options for mediation.

**Use of spectra as a transient detector substitute[2].** Emerging technologies are increasingly developing this approach for pitch, with Melodyne being followed by Pro Tools v.8 shortly. This could be extended to spectral fingerprinting to resolve audio such as rapid successive crashes, which show neither clear transients or contain pitched information.

### Visualization

A crude animation will endeavor to illustrate a small portion of this specification, not as a proposed GUI, but merely as elucidation of the above, focusing on the hybrid grid and click.



Movie 26: This animation implies one aspect of the feature set described: the use of a multi-touch interface to set discrete areas of a take to a given quantize value.

## 4 Conclusions

All of the systems appraised exhibit strengths and weaknesses, many of which have not been highlighted here; the focus for this study has been qualities deemed relevant to comparison with each other. There are many more systems that have not been considered here, but which if studied similarly would contribute to an even more comprehensive understanding of current workflows. Slicing techniques mostly require extensive mediation to compensate for tempo drop induced gaps, but in fairness can work well for tempo increases. Stretching with a suitable algorithm within the manufacturer's specified range can produce the most transparent effect, but that range is all too easily over-stepped in 'normal' musical procedures, leading to artifacts. Many principal DAWs tend to implement 3rd party stretch algorithms, and so given that these will produce the same aesthetic sonic quality, it is indeed the workflow that will lead to a user 'preference', in addition of course (and rightly often subservient) to perception of this audio quality. Three of those studied here (Logic, Pro Tools and Melodyne) do offer proprietary algorithms, and naturally there is validity in comparing those to the more ubiquitous third party ones in another work. All systems are in continual development and so where one might appear to be superior at a given task, the others may soon supersede.

[2] Transient detection processes typically operate in the frequency domain anyway.

The opportunity still exists for a definitive system, and here also lies future work!

## 5 Acknowledgments

Thanks are due to Jerry Fleming and Sandor Jozsa who produced recordings to assist this work. The purpose of these recording was to evaluate time-stretch performance differences between normal and 'high' sample rates, but ultimately this was deemed beyond the scope of this paper and may form the subject of a future study.

## References

Bachmann, C., Bischoff, H., Grossmann, B., Pfeifer, S. and Schlomberg, C. 2006. *Cubase 4 Operation Manual*. Steinberg Media Technologies.

Guitton, Jean-Stéphane. (2008) *Time Stretching & Pitch Shifting: Comparison Part I*. Available at: http://plugin-pitch-shifter-time-stretcher.en.audiofanzine.com/instruction/print_dossier,idossier,10.html [Online]. [Accessed 11 December 2008]

Han, Jeff Y. (2006) *Multi-Touch Interaction Research*. Available at: http://www.cs.nyu.edu/~jhan/ftirtouch/ [Online]. [Accessed 12 December 2008].

Hoenig, U. G. and Neubäcker, P. 2005. *Melodyne cre8/studio user manual Version 3.x*. München: Celemony Software GmbH.

http://www.izotope.com/press/9.26.2006_iZotope_Licenses_Radius_Technology_to_Digidesign.html (2008) [Online]. [Accessed 12 December 2008].

http://knowledgebase.steinberg.de/259_1.html (2008) [Online]. [Accessed 11 December 2008].

http://www.propellerheads.se/press/index.cfm?fuseaction=get_press_release&prID=23 (2004) [Online]. [Accessed: 11 November 2008]

Pye, D. (2008) Logic Pro Help 7th January 2008 [Online]. Available at: http://logicprohelp.com/viewtopic.php?t=19457 [Online]. [Accessed: 6 November 2008].

http://zenph.com/sept25.html (2008) [Online]. [Accessed 12 December 2008].

http://www.zplane.de/index.php?page=description-elastique (2008) [Online]. [Accessed 11 December 2008].

## Bibliography

Apple Inc.. 2007. *Logic Pro 8 User Manual*. Cupertino, CA: Apple Inc..

Carlson, L., Nordmark, A. and Wiklander, R. 2001. *Recycle Operation Manual*. Stockholm: Propellerhead.

DeSantis, D., Haywood, K., Knudsen, R., Behles, G.,Rang, J., Henke, R. and Slama, T. 2007. *Ableton Live 7: Reference Manual*. Berlin: Ableton AG.

Digidesign. 2007. *Pro Tools 7.4 Reference Guide*. Iver Heath, UK: Digidesign.