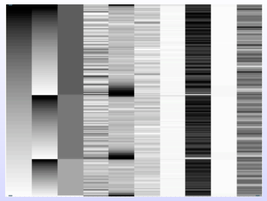
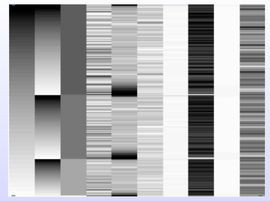


Beyond Concatenation: Some Ideas for the Creative Use of Corpus-based Sonic Material

Thomas M. Stoll

Hiller Music Studios, University at Buffalo, Buffalo, NY, USA



Abstract

Concatenative synthesis and audio mosaicing are two techniques that utilize large corpora of sounds to facilitate data-driven sampling. This paper reviews some of the organizational strategies implicitly applied in concert with these techniques and proposes several ideas of how to deal with units in ways that incorporate compositional objectives. Some general ideas on how to extend and make adaptable the underlying frameworks of corpus-based synthesis are presented.

Grains, Micromontage, Mosaicing, and Concatenation

In concert with an atomistic concept of audio proposed by Dennis Gabor, granular approaches to creating sound materials have taken diverse forms. Micromontage, audio mosaicing, and concatenative synthesis all rely on various interpretations of the particulate concept. The author has used CataRT to explore and learn about concatenative synthesis. The visual examples depicting corpora were made with a custom version of CataRT adapted by the author. This adapted corpus-based system is currently set up to deal with audio segments as short as tens of milliseconds and as long as several seconds. In most cases, the efficacy of representing a unitary segment of sonic material diminishes as more information (more duration) is included, but this is ultimately left to the user to determine.

Any composition, especially that of musique concrte, can be thought of as employing certain tangible materials. Thus, any composition's set of source materials may be thought of as a corpus of sounds from which the composer selects individual members for deployment at a particular point in time for a particular compositional purpose. This process, differing greatly from a composer using an abstract representation (a score) of ideal sound sources (performers), highlights the visceral process of creating music out of the sounds themselves. Mosaicing and concatenative sound synthesis rely on data associated with corpora of sounds for organization and deployment. These techniques are explicitly corpus-based, commonly relying on large pools of potential sound units (segments) to be matched to targeted specifications. In order for corpus-based searching to perform efficiently, matching is based on sets of derived features from windowed audio.

The concept of a unit is implicit in both audio mosaicing and concatenative synthesis, but the unit plays a different role in each. This paper examines prevalent models for conceptualizing units within a corpus, introduces several ways that units can be organized such that order and other parameters derived from an input sound are used as primary organizing factors, and discusses additional concepts and strategies for organizing corpora of sounds that go beyond metaphors of space.

Organizational Strategies

Implementations of corpus-based sampling draw their organizational strategies from two dominant models: that of a predefined trajectory through a (multidimensional) target space and that which concerns free navigation of parameter space. Both focus on good local choice based on the minimization of perceptual distance measures.

The target trajectory stratagem, while having the potential to reveal interesting emergent properties of differentials that arise from changes of target corpus, lacks an immediate facility for surprise, as the search process is bound by reference to the source. There is a certain type of freedom in the navigational mode (shown in Figure 1a). It assumes that connections can be intuited from proximity between two units on a graph and observable similarity of those sounds. Overall, there is a tradeoff between control that captures some sense of the control sound's nonlinearity and navigational freedom that relies heavily on proximity.

Perhaps these assertions result from underestimation of the potential for a freer kind of audio mosaicing, where the most extrapolated yet still perceptually referential (to the control sound) results are returned. Put another way, it may be possible to imagine a kind of constructive distortion on the unit level by employing an incomplete, broken, or chaotic search method. The combination of the notion of order borrowed from mosaicing and the freedom of proximate association borrowed from concatenative synthesis makes way for hybrid variations of these schemes.

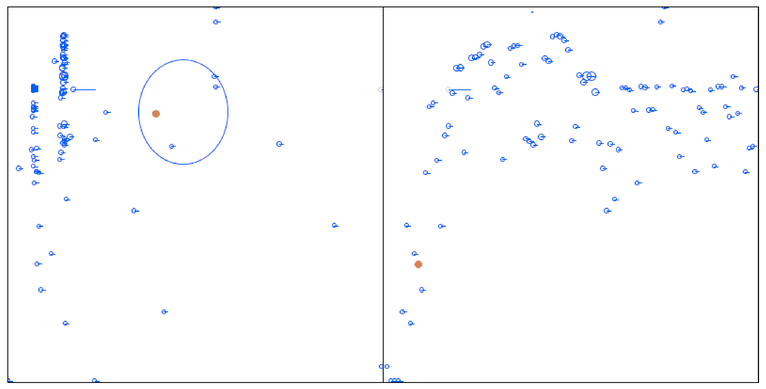


Figure 1: (a) Free navigation through points, in this case x = pitch and y = periodicity. The large oval is a search radius. In (b) x = order, allowing for the order within the source sound to be preserved.

Beyond Navigation: Order As A Parameter, Anchor Units, and Networks

The use of orderas derived from the inputcomprises a simple refinement of the system that emphasizes continuous gradation. In its simplest form, the use of order would render the process equivalent to that of making cuts in an audio file in a sound editor. The addition of a parameter to the vertical axis (assuming that time is shown on the horizontal axis) creates a secondary hierarchy for exploitation by a search process (see Figure 1b). The search algorithm could occasionally jump to a unit with a close match in the secondary parameter, thus disrupting the predefined order of segments. The algorithm would have some criteria for deciding under what conditions to jump, and would resemble something like nonlinear distortion, only on the unit level rather than the sample level.

A unit may be designated to be relatively more important than others based on some measure, (e.g. loudness or duration), with respect to order within the source sound file, or by a hand-selection process. This hierarchical classification makes possible the elevation of some units to the status of anchor units (see Figure 2). These anchors may be fixed as reference points and could allow for the redistribution of non-anchor units. The user would then be free to reposition these free units in relation to the fixed units according to some secondary parameter(s) or by hand. An ordering principle may be implemented, but does not have to be.

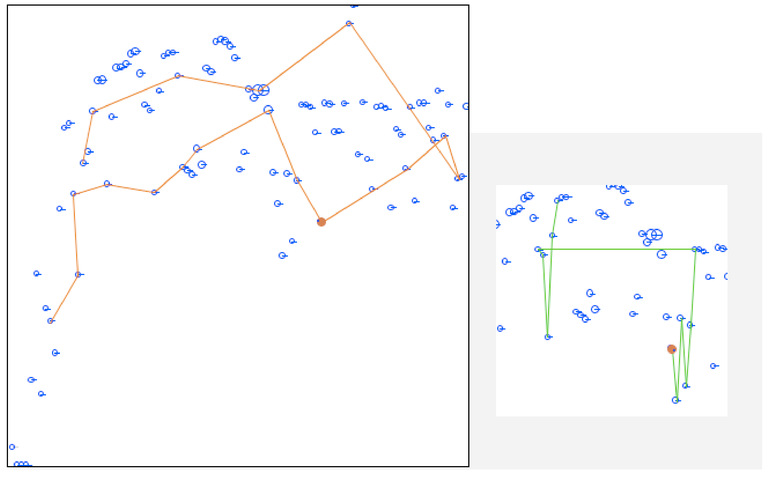


Figure 2: (a) A network of anchor units (without redistribution). (b) Navigation could follow some rules that respect the ordered dimension some of the time, only sporadically moving to a nearby unit in the other dimension.

Classification Networks

The above grouping schemes are primarily based on respecting the individuality of each unit. An alternative approach involves a grouping stage, where a series of binary classifications would assign each unit in a corpus a distinct fingerprint, with no limit to the number of units that might share a certain fingerprint. Search methods operating on a collection of fingerprinted units would be able to make selections among sets or sets of sets. This approach has been used by the author with some success, and is important because it can be used to selectively narrow down the units that might be presented to a distance or order-based selection algorithm. It may be easier to visualize this procedure as a graph of two such classifications (see Figure 3).

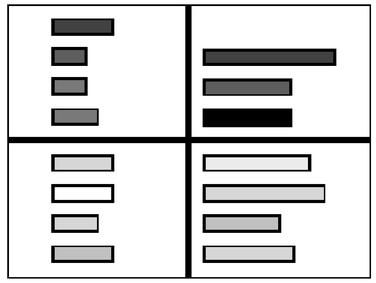


Figure 3: Units are placed into distinct categories according to some simple rules. Amplitude is shown by shading and duration is shown by segment length.

Other Concerns

It might seem to be a step backward to remove or de-emphasize the search-oriented nature of concatenative synthesis, but by focusing on organizational potential rather than search, new and inspiring possibilities arise, including the potential to redefine the corpus as a dynamic, developing entity within a musical composition or improvisation.

1. Dynamic or Recursive Segmentation: Any unit that can be isolated as a discrete entity according to one principle may be redivided according to some other principle. A stream that is divided according to detected changes in amplitude values may be redivided according to detected changes in pitch or some other pattern. In addition to the two order-based models outlined above, one can envision a more general approach to segmentation with the possibility for dynamic, recursive, or other mutable procedures based on the recognition that segmentation can be simultaneously performed in different ways and on more than one time scale.

2. Static and Dynamic Formal Classification: The segmentation process can be extended to incorporate formal aspects, either set by the user or detected in the course of the tracking process. A criterion may be invented for a piece of music that identifies a particular formal element to be used throughout a piece. A dynamic method could attempt to analyze incoming units' likely formal classification: note, partial phrase, phrase, etc. In any case, a hierarchical classification scheme may be desired to mark levels of membership. Larger units would contain smaller units, and units could be members of more than one group. Methods to control the administration of such a scheme could be made accessible to any part of a distributed, modular system.

3. Proximity and Multiplicity of Representation: One might draw the conclusion that proximity in a linear sense, even if considered across many dimensions, has no place in a system not based on graphically scratching/scanning through points. However, once a unit has been selected from an order-based or n-ary classification scheme, additional units may exist similar to that unit. A search by proximity could then furnish the final decision as to the exact unit to deploy. This two-stage retrieval, where two separate, compatible strategies are used, reveals the power of a simultaneous multiplicity of representation. By decoupling search into categorical and proximate representations, non-overlapping decisions are easier to imagine and execute. The combination of formal designation and search allows for interaction and real time musicality.

4. Mutability: It should be fairly straightforward to extend the corpus model to include a definition of the unit as a mutable object, with full facility to derive, modify, and recombine units. Furthermore, the processes of organization themselves should be open to modification and adaptation, either automatically or interactively.