

**Perspective, Texture, Density, Motion:**

## **Aesthetics and the Art of Audio Field Recording**

**By Steven M. Miller**

Paper presented in the symposium *Music and ecologies of sound. Theoretical and practical projects for a listening of the world*, University Paris 8, May 2013

### Introduction

When approaching audio field recording (often referred to as ‘phonography’) as a sound art practice, inevitable questions of aesthetics surface. The common attitude often expressed or implied runs something along the lines of “What is the artistic element of something so seemingly ‘passive’ an activity as pointing a mic and pushing the record button?” (reminiscent of similar historical attitudes about photography as a fine art). Without addressing the implicit assumptions of transparency of mediating technology, or getting into the hands-on techniques involved (both worthy subjects and amply covered elsewhere), I’ll focus instead on the aesthetic qualities of the recordings themselves that reward focused, intensive (i.e., active and engaged) listening.

When listening to sounds we can choose to focus on the perceptual qualities of the sounds, or on the meanings that the sounds convey. Of course, the situation is usually much more fluid than this simple ‘this or that’ dichotomy; we usually transition rather seamlessly between attentional strategies to shift focus between any number of aspects of the sonic experience. Perceptual qualities are only the most obvious layer of sounds, but are also often the most easily ignored. Other qualities may include semantic, syntactic, contextual, and/or symbolic meanings conveyed by the sounds. In terms of a communicational experience of sound, these ‘other’ qualities tend to play a more prominent role. This post, however, in considering the qualitative aesthetic aspects, will focus on the perceptual qualities of the sounds themselves.

The four primary qualities or characteristics I listen for in making engaging field

recordings are: *perspective*, *texture*, *density*, and *motion*. In evaluating the degree of aesthetic interest, these are the aspects I find myself listening to (or for). One note worth making here: this list is not exhaustive, nor are the individual characteristics independent and/or mutually exclusive. Combinations of various sorts are not only possible but quite common, and the categories often exhibit a high degree of interdependence.

### *Perspective*

Perspective is often understood as referring to the idea of a vantage point, either in the visual sense or – through metaphor – in the conceptual sense. Definitions from the Oxford dictionaries online include, “A particular attitude toward or way of regarding something; a point of view” and “An apparent spatial distribution in perceived sound”.

([http://oxforddictionaries.com/definition/american\\_english/perspective](http://oxforddictionaries.com/definition/american_english/perspective)) It is this second definition on which I’ll focus.

Aural perspective is typically achieved/experienced as a function of the following:

*Relative proximity* – distance from listener to sound source(s)

Is it, or are they, close or distant? If multiple sounds are audible, is there a combination of close and distant sources?

*Panorama* – left-to-right distribution

Is the sound field wide or narrow, concentrated or distributed? Is the perceived sound field symmetrical or asymmetrical?

*Foreground/Background* – relative focus of attention

While often closely related to relative proximity, other aspects such as relative loudness, timbre, articulation, and rates/degrees of activity and change can also influence the balance between sounds experienced as primary focus points versus those experienced as background.

### *Texture*

While texture can be understood in a number of ways, here I’m referring to

perceived aspects of surface quality. Again, Oxford online defines it as, “The feel, appearance, or consistency of a surface or a substance” and, “The quality created by the combination of the different elements in a work of music or literature”. ([http://oxforddictionaries.com/definition/american\\_english/texture](http://oxforddictionaries.com/definition/american_english/texture))

In relation to sound(s), then, texture can be thought of as resulting from the intersection of the following parameters:

*Timbre(s)* – ‘tone color,’ spectrum, and envelope

Often glossed simply as ‘sound quality,’ timbre is a notoriously slippery term to define. Most contemporary definitions deal more with what it is not than what it is (i.e., everything other than pitch, loudness, and duration (and location, I might add)), but it can be summarized as the quality of sound that allows us to distinguish different instrumental families, for example, or different sound sources from each other. It is often characterized as stemming from the presence/absence and relative strengths of different simultaneous frequency components and the averaged changes in amplitude over the span of a single sound event (i.e., attack, decay, etc.). In general, it can be linked, as well, with ideas of sonic morphology.

We often will describe timbral qualities with terminology borrowed from other senses: we say something sounds ‘hollow’ or ‘metallic’ or ‘thin’ or ‘bright.’

*Articulation* – characteristics of ‘onset’ and ‘offset’

In musical performance, articulation refers to the manner of playing notes, for example smoothly bowing, crisply striking, or plucking the strings of a violin are all different types of articulation. This can be generalized more broadly as the manner in which the resonating body is energized; how is energy transferred from one object into another. This in turn affects timbre-related aspects such as spectrum, morphology, and envelope, as well as peak amplitude and (to a lesser degree) pitch.

We will often describe a sound’s articulation as sharp or dull, crisp or muted, or by analogy to musical instruments, i.e. percussive or smooth. In musical terminology we might say staccato or legato, etc.

*Surface characteristic(s)* – a meta-level of timbre and articulation

In a single complexly changing sustained sound or aggregates of multiple sounds, the overall net effect of numerous individual changes in texture can be characterized as, for example, rough or smooth, and the surface of these changes – how they progress over time – as linear, angular or curved.

### *Density*

Density of sound can be thought of as a function of multitude and proximity (in time, in location, in any other perceivable parameter). Oxford defines it as, “The degree of compactness of a substance” and, “The quantity of people or things in a given area or space”

([http://oxforddictionaries.com/definition/american\\_english/density](http://oxforddictionaries.com/definition/american_english/density)) – which translates in sonic terms to how ‘spread out’ the sound is, on some given axis, and the relationships among/between the various elements.

### *Temporal* (Horizontal)

Density of elements in time: rate of activity and/or rate of change. In music this equates to pulse, rhythm, subdivision, and tempo.

### *Textural* (Vertical)

The number of simultaneous discrete events/layers is the textural density of sound(s). In music, this would relate to harmony, voicing, and orchestration.

### *Foreground/Background*

Though introduced under ‘perspective’ above, here the question of foreground/background deals with the relative amount of focal attention given to individual elements or layers of elements given their density. A common occurrence here is the phenomenon of ‘masking’ whereby one sound/layer obscures another.

### *Hi-fi vs. Lo-fi*

In “The Tuning of the World” Murray Schafer introduces the terms hi-fi

and lo-fi as describing soundscapes in which individual sounds can be heard with clarity (hi-fi) or those in which individual elements are lost amongst the multitudes of sounds (lo-fi). One obvious aspect here is the relative density of these two types of soundscapes.

### *Motion*

According to Oxford online, motion is “The action or process of moving or being moved”. ([http://oxforddictionaries.com/definition/american\\_english/motion](http://oxforddictionaries.com/definition/american_english/motion))

Sound sources appear to move, or we perceive motion, due to a number of factors, to include actual motion of the sound source(s) and/or motion of the listener/recorder. In stereo recordings this motion (or perceived motion) is two-dimensional: left-to-right (horizontal panorama) and front-to-back (horizontal proximity). Surround sound (multi-channel) recordings expand the front-to-back dimension to envelope the listener. Motion can be implied, however, where none in actuality exists, for example through rapid alternation of individual stationary sound sources. Such auditory illusions have visual corollaries.

There are four primary categories of motion in the soundscape or audio recordings.

*Static* – no apparent motion

*Animated* – implied motion resulting from interaction of a number of individually motionless sound sources

*Object(s) in motion* – sound source(s) moving

*Subject in motion* – listener moving

### *Conclusion*

There are numerous levels on which a field recording may be judged aesthetically; I've focused here on the 4 primary parameters that I evaluate when listening to or making recordings. I have intentionally not specified 'good/bad' values or ranges for these parameters. My experience is not so much that more or less of any one of these categories is necessarily better or worse; rather, to the extent that you pay attention to them and work with them to get the result you are after, your recordings will be more interesting and aesthetically engaging

to listen to.

### *Listening Examples*

The following listening examples are all drawn from this blog – not because I think my recordings are necessarily better than others; rather simply because they're the ones with which I'm familiar from both the 'making' and 'listening' sides of the experience. They are presented in no particular order, along with short explanatory notes linking them to the discussion above.

#### Listening Example 1

Frog pond – Flower Mountain, Bali

<http://sightssoundswords.wordpress.com/2012/01/09/frog-pond-2/>

*Perspective* – a foreground mixture of near-field and slightly more distant frogs, with far-distant insects in the background; widely distributed and fairly symmetrical.

*Texture* – somewhat granular, consisting of a number of relatively short sounds that are smoothly articulated.

*Density* – reasonably low density; each sound is clearly audible, even the distant and quiet insects; temporal density fluctuates over the course of the recording, while textural density remains very low throughout.

*Motion* – animated; mostly side-to-side with lesser front-to-back.

*Comments* – recorded right at the edge (and about 4–6" off the ground) of a small frog pond that is about 2–3 feet across from front to back and about 8–10 feet wide; surrounded by trees and low bushes; near the village of Payangan, Bali.

#### *Listening Example 2*

Street-side café – Den Haag, Netherlands

<http://sightssoundswords.wordpress.com/2012/04/04/den-haag->

*sidewalk-cafe/*

*Perspective* – background of distant traffic barely intrudes on a foreground of pedestrian, bicycle, and motor scooter traffic; widely distributed panorama enhanced by the considerable lateral motion which is fairly symmetrical, on average; proximity of foreground elements ebbs and flows with their continuous approach and retreat; some background elements are farther off, including seagulls and doors opening/closing.

*Texture* – smooth distant hum with percussive foreground foot-falls of pedestrians and rough surface of passing bicycles and scooters.

*Density* – medium temporal and textural density; despite the urban context and presence of background traffic noise, the overall effect is relatively hi-fi, with each fore- and mid-ground sound source clearly audible.

*Motion* – predominantly objects in motion, dominated by side-to-side panoramic movement; no discernable front-to-back motion other than one car which backs away from the listener's position; a few individual background elements are stationary.

*Comments* – recorded sitting at a table, on a cold clear morning, just outside the door of a café situated on a cobblestone street that is closed to automobile traffic; the street is lined with brick and stone buildings, mostly commercial establishments, and somewhat low-rise architecture contributing to the clear presence of the foreground and passing elements.

### *Listening Example 3*

Old town plaza – Corfu, Greece

<http://sightssoundswords.wordpress.com/2012/02/14/corfu-town-plaza/>

*Perspective* – mixture of near-field and more distant layers; symmetrical and evenly distributed wide panorama.

*Texture* – smooth with spikes and bumps; dense background wall forms a

smooth backdrop against which the foreground elements stand out.

*Density* – the strongest and most obvious initial impression is of extremely high density, both temporal and textural, which is fairly uniform and consistent throughout the recording; upon closer listening, despite the extreme density and resulting overall lo-fi impression, there is still a clearly discernable set of foreground sounds which distinguish themselves from the background wall of sound.

*Motion* – a combination of static, animated, and side-to-side motion; background elements fuse into a static sound field, with occasional animated elements; foreground elements exhibit both panoramic and animated motion.

*Comments* – recorded mid-afternoon while sitting at an outdoor café on the plaza, the sheer density is striking, though the seemingly lo-fi character is mitigated by an active and clearly audible foreground.

#### *Listening Example 4*

Riverside along the Seine – Paris, France

<http://sightssoundswords.wordpress.com/2011/11/26/pont-saint-louis/>

*Perspective* – two distinct layers, one extremely close-up and the other quite distant; in an unusual twist, the closer layer is not always the focal point/foreground, and the distant layer is not always background.

*Texture* – a somewhat angular layer over a more solid and thick layer; fluid timbres contrasted with metal and machine.

*Density* – both layers have a fairly high temporal density; the distant layer is texturally dense while the close-up layer is texturally sparse.

*Motion* – close-up layer is quite active with continual and at times extreme motion (enhanced by mic proximity), while distant layer is static.

*Comments* – recorded with the mic suspended about 2–3” above the water surface at the edge of the river on a stone pier, allowing the water

to flow up to, under, and behind the mic; exploits the interplay of perspective, density, and motion resulting in a paradoxical situation where the active, highly mobile, close-up layer often recedes to the perceptual background while the static-motion distant layer comes to the foreground, largely due to the repetition and consistency of the water versus the continually evolving bell and city sounds.

### *Listening Example 5*

Evening sounds – Bandung, West Java

<http://sightssoundswords.wordpress.com/2011/12/29/bandung-evening/>

*Perspective* – mixture of mid-field and distant elements; wide panorama that is symmetrically arrayed.

*Texture* – smooth and even, punctuated with shorter, sharper, and rougher elements; the mechanical timbres are offset with a number of human and animal sounds; shorter and dynamic sounds are contrasted with sustaining and slowly evolving ones.

*Density* – medium temporal density, with a textural density that varies from medium to high (as the calls-to-prayer begin to predominate).

*Motion* – a combination of animated and objects-in-motion; the side-to-side motion is complimented by a fair amount of front-to-back motion implied in the voices of the itinerant vendors selling saté.

*Comments* – perched on the 3rd-floor balcony of my hotel room, overlooking a low, primarily residential hillside, with small lanes traversed by motorcycles, cars, bicycles, and pushcart food vendors; the evening call to prayer starts as a sparse background element then soon develops into a thickly dense and evolving focal point.

### *Listening Example 6*

Insect drone – Norman, Oklahoma

<http://sightssoundswords.wordpress.com/2012/11/20/insect-drone/>

*Perspective* – mid-field proximity; evenly distributed and symmetrical wide panorama.

*Texture* – grainy and rough; shrill timbres.

*Density* – depending on how you consider it, either very low or very high density, both temporally and texturally: the numerous individual elements perceptually fuse into a single aggregate; considered individually there are a high number of individual elements all overlapping, while considered as a single aggregate there's not much change over time and only one layer.

*Motion* – primarily static with touches of animated.

*Comments* – a great example of a soundfield that on the surface seems very singular and undifferentiated yet on closer listening reveals numerous individual events with a surprisingly dynamic internal structure.

#### *Listening Example 7*

YSTCM windows – Singapore

<http://sightssoundwords.wordpress.com/2012/10/06/ystcm-windows-1/>

*Perspective* – somewhat anomalous, given the recording technique of contact mics attached to large plate-glass windows; despite this, the illusion persists of some elements close-up with others more distant.

*Texture* – mostly smooth and evenly flowing.

*Density* – texturally fairly dense while temporally less so; most individual elements sustain for long periods, yet a number of them have significant internal changes of timbre and/or amplitude.

*Motion* – the surface perception of static motion is undermined by the constant change and evolution of most individual elements, leading to an illusion of animated motion.

*Comments* – like the previous example, this recording yields more and more dynamic detail upon close listening.

### *Listening Example 8*

Ghent-to-Brussels train ride – Belgium

<http://sightssoundwords.wordpress.com/2012/07/03/ghent-to-brussels-4/>

*Perspective* – primarily close and mid-field, interior ambience, audibly small-sized space; asymmetrical panorama with conversational voices predominantly to one side with mechanical train sounds more evenly distributed; shallow sense of front-to-back depth.

*Texture* – clear and clean foreground layer of conversation, announcements, and signals over a smoothly rhythmic background layer of sustaining train sounds.

*Density* – episodic alternation between low and medium textural density (as the train is either stopped or in motion), coupled with a low temporal density foreground layer over the recurring sustained background.

*Motion* – despite the motion of the train, the soundfield is primarily static.

*Comments* – recorded with the mic laying on the small table in front of my seat, padded from the direct vibrations by a jacket; a good example of lots of absolute motion (the train and everything in it) leading to primarily static perception (very little relative motion).

### *Listening Example 9*

Chimayo soundscape – Chimayo, New Mexico

<http://sightssoundwords.wordpress.com/2012/07/08/chimayo-soundscape-4/>

*Perspective* – mid-field and distant elements; somewhat asymmetrical medium-wide panorama: lower-pitched water drips off to the left with a

barely audible slap-back echo from a concrete retaining wall off to the right, higher-pitched drips in the center.

*Texture* – rough and choppy water over smooth but grainy insects.

*Density* – medium temporal and textural density with the water drips providing most of the temporal aspect and the insects, birds, and plane providing the textural elements.

*Motion* – mixture of static and animated.

*Comments* – nice (though very subtle) example of naturally occurring delay/echo due to positioning the mic to catch the sound off to one side and its reflection off the concrete wall off to the other.

### *Listening Example 10*

Soi soundwalk – Bangkok, Thailand

<http://sightssoundwords.wordpress.com/2012/09/06/soi-kasem-san-2/>

*Perspective* – varying perspective as the mic moves in relation both to stationary and moving sound sources; asymmetrical panorama due to walking along the left edge of the road.

*Texture* – complex and ever-changing.

*Density* – reasonably high temporal and textural density due to urban setting.

*Motion* – combination of object-in-motion and subject-in-motion as I walk the length of the soi (small side street) from the main road to the klong (canal).

*Comments* – a mid-morning walk along a busy soi (side street) from a busy main thoroughfare towards a klong (canal) in the heart of Bangkok; the walk ends in front of the Jim Thompson House.

### *Listening Example 11*

Along the klong – Bangkok, Thailand

<http://sightssoundwords.wordpress.com/2012/09/26/ban-krua/>

*Perspective* – varying perspective as the mic moves in relation both to stationary and moving sound sources; asymmetrical panorama due to walking first along the left edge of the canal then along the right.

*Texture* – complex and changing.

*Density* – alternating between relatively low-density and periods of high density in both the temporal and textural dimensions.

*Motion* – combination of object-in-motion and subject-in-motion as I walk along the klong (canal) then cross over a bridge and continue walking along the other side.

*Comments* – a mid-morning walk through a typical neighborhood along the canals of Bangkok, with a variety of sounds, moods, and textures, exhibiting the possible diversities often hidden in urban settings. There are two sections of technical flaws within this recording consisting of less-than-transparent dynamic range compression/limiting; despite this, I very much like the overall feel of this recording.

### *Listening Example 12*

Jim's soi-side klong – Bangkok, Thailand

<http://sightssoundwords.wordpress.com/2012/09/21/klong-maha-nag-3/>

*Perspective* – mixture of medium and distant with neither predominating.

*Texture* – choppy and rough water punctuated by bird and human vocalizations, a distant motorcycle, and other assorted more percussive sounds.

*Density* – fairly consistent and homogenous sense of layering with little standing out as obvious foreground material other than the few occurrences of human speech.

*Motion* – combination of static, animated and object-in-motion.

*Comments* – mic at chest-level, standing near the edge of the klong (canal) at the end of the soi (side street), just in front of the Jim Thompson House.