

The Art of Recording

The Artistic Elements of Recording Production

As will be remembered, sound in audio recording is in three states: physical dimensions, perceived parameters, and artistic elements.

The artistic elements are the resources of the recordist for artistic expression. The perceived parameters translate into the artistic elements:

- (1) pitch becomes pitch levels and relationships;
- (2) loudness becomes dynamic levels and relationships;
- (3) duration becomes rhythmic patterns and rate of activity;
- (4) timbre becomes sound sources and sound quality;
- (5) space becomes spatial properties

The audio production process provides the resources for considerable variation, and the very refined control of ALL of the artistic elements of sound. This allows all of the artistic elements of sound to be accurately and precisely controlled through many states of variation, in ways that were possible with ONLY pitch on traditional musical instruments.

THE STATES OF SOUND IN AUDIO RECORDING

| PHYSICAL DIMENSIONS | PERCEIVED PARAMETERS | ARTISTIC ELEMENTS |
|---|--|---|
| Acoustic State) | (Psychoacoustic Conception) | (Resources for Artistic Expression) |
| Frequency | Pitch | Pitch Levels & Relationships melodic lines, chords, register, range, tonal organization, textural density, pitch areas, vibrato |
| Amplitude | Loudness | Dynamic Levels & Relationships dynamic contour, emphasis accent, tremolo, balance (dynamic relationships of sound sources) |
| Time | Duration time perception | Rhythmic Patterns & Rates of Activities tempo, time, patterns of durations |
| Timbre (comprised of physical components: dynamic envelope, spectrum, & spectral envelope) | Timbre (perceived as overall quality; defined by the definition of fundamental frequency and by the analysis of the Physical Dimensions) | Sound Sources and Sound Quality sound sources, groupings of sound sources, instrumentation, performance intensity, texture (quality of the overall sound), performance techniques |
| Space (comprised of physical components created by the interaction of the sound source & the environment, and their relationship to a receptor) | Space (perception of the sound source as it interacts with the environment, and perception of the physical relationship of the sound source to the listener) | Spatial Properties lateral (stereo) location, imaging, distance, moving sources, depth of sound stage, phantom images, environmental characteristics, space within space |

Pitch Levels and Relationships

Relationships of pitch levels contain most of the pertinent information in a piece of music. The artistic message of most of the music of Western heritage is communicated (to a large extent) by pitch relationships. The listener has been trained, by the music heard throughout their life, to focus on this element to obtain the most significant musical information. The other artistic elements often support pitch patterns and relationships.

Pitch is the most precisely controlled artistic element in traditional music. The use of pitch relationships and pitch levels in music is more sophisticated than the use of the other artistic elements. Complex relationships of pitch patterns and levels are quite common in music.

Information about the artistic element of pitch levels and relationships will be related to

- (1) the relative dominance of certain pitch levels,
- (2) the relative registral placement of pitch levels and patterns, or
- (3) pitch relationships: patterns of successive intervals, relationships of those patterns, and relationships of simultaneous intervals.

The artistic element of pitch levels and relationships is broken into the component parts: melodic lines, chords, tonal organization, register, range, textural density, pitch areas, and tonal speech inflection.

A series of successive, related pitches creates melodic lines. Melodic lines are perceived as a sequence of intervals that appear in a specific ordering, and that have rhythmic characteristics. The melodic line is often the primary carrier of the artistic message of a piece of music.

The ordering of intervals, coupled with or independent from rhythm, creates patterns. Pattern perception is central to how humans perceive objects and events. These basic principles relate to all of the components of the artistic elements. Melodic lines are organized by patterns of intervals (short melodic ideas, or motives), supported by corresponding rhythmic patterns. The complexity of the patterns, the ways in which the patterns are repeated, and the ways in which the patterns are modified provide the melodic line with its unique character.

Two or more simultaneously sounding pitches create chords. In much of our music, these chords are based on superimposing, or stacking, the intervals of a third (intervals containing three and four semitones, most commonly). Chords comprised of three pitches, combining two intervals of a third, are called triads. The continued stacking of thirds results in seventh, ninth, eleventh, and thirteenth chords.

The movement from one chord to another, or harmonic progression, is the most stylized of all the components of the artistic elements. Harmonic progression is the pattern created by successive chords, as based on the lowest note (the root) of the triads (or more complex chords). These patterns of chord progressions have become established as having general principles that occur consistently in certain types of music. Certain types of music will have stylized chord progressions (progressions that occur most frequently), other types of music will have quite different movement between chords, and perhaps emphasize more complex chord types. The patterns of the harmonic progression create harmony.

Harmony is one of primary components that supports the melodic line. Pitches of the melody are reinforced by the chords in the harmonic progression. The speed and direction of the melodic line is often supported by the speed at which chords are changed, and the patterns created by the changing chords: harmonic rhythm.

The expectations of harmonic progression create a sequence of chords which will present areas of tension and areas of repose within the musical composition. The tendencies of harmonic motion do much to shape the momentum of a piece of music, and can greatly enhance the character of the melodic line and musical message. Performers utilize the psychological tendencies of harmonic progression; exploiting its directional and dramatic tendencies. The expectations of harmonic movement and the psychological characteristics of harmonic progression have become important aspects of musical expression and musical performance.

The melodic and harmonic pitch materials are related through tonal organization. Certain pitch materials are emphasized over others, in varying degrees, in nearly all music. This emphasis creates systems of tonal organization in which a hierarchy of pitch levels exist. A hierarchy will most often place one pitch in a predominant role, with all other pitches having functions of varying importance, in supporting the primary pitch. The primary pitch, or tonal center, becomes a reference level, to which all pitch material is related, and around which pitch patterns are organized.

Many tonal organization systems exist. These systems tend to vary significantly by cultures, with most cultures using several different, but related systems. The "major" and "minor" tonal organization systems of Western music are examples of different, but related systems; as are the "whole-tone" and "pentatonic" systems of Eastern Asia. The reader should consult appropriate music theory texts for more detailed information on tonal organization, as necessary.

Certain components of pitch levels and relationships have become more prominent in musical contexts (and other areas of audio) because of the artistic treatment pitch relationships have received in music recordings. The components of range, register, textural density, and pitch area are able to be more closely controlled in recorded music, than in live (un-amplified) performance. These components are more important in recorded music, because they are precisely controllable by the technology, and they have been controlled to support and enhance the musical material.

Range is the span of pitches of a sound source (or of any instrument or voice). Range is the area of pitches that encompasses the highest note possible (or present in a certain musical example or context) to the lowest note possible (or present) in a particular sound source.

A register is a portion of a sound source's range. A register will have a unique character (such as a unique timbre, or some other determining factor) that will differentiate it from all other areas of the same range. It is a small area within the source's range, that is unique in some way. Ranges are often divided into many registers; registers may encompass a very small group of successive pitches, up to a considerable portion of the source's range.

A pitch area is a portion of any range (or of a register) that may or may not exhibit characteristics that are unique from other areas. Instead, it is a defined area between an upper and a lower pitch-level, in which a specific activity or sound exists.

Textural density is the relative amount and registral placement of simultaneously sounding pitch material, throughout the hearing range or within a specific pitch area. It is the amount and placement of pitch material in the composite musical texture (the overall sound of the piece of music) in relation to defined boundaries.

With textural density, sound sources are assigned (or perceived as being within) a certain pitch area, within the entire audible range (or range used within a certain piece of music). Thus, certain pitch areas will have more activity than other pitch areas; certain sound sources will be present only in certain pitch areas, and other sources present only in other pitch areas; some sources may share pitch areas, and cause more activity to be present in those portions of the range; some pitch areas may be void of activity. Many possible variations exist.

Textural density is a component of pitch-level relationships, that is directly related to traditional concerns of orchestration. Textural density is a much more specific concern in recorded music because it is controllable in very fine increments. Traditional orchestration was concerned, basically, with the selection of instruments, and with the placement of the musical parts (performed by the assigned instruments) against one another.

With the controls of signal processing (especially equalization), sound synthesis and multi-track recording, the registral placement of sound sources and their interaction with the other sound sources take on many more dimensions. Each sound source occupies a pitch area; the acoustic energy within the pitch

area of a timbre's spectrum is distributed in ways that are unique to each sound source. The spectrum of each sound source is an individual textural density, and the textural density of the overall program (or musical texture) is the composite of all of the simultaneous pitch information from all sound sources.

Sound sources, and musical ideas, are often delineated by the pitch area they occupy within the composite textural density. Sound sources are more easily perceived as being separate entities and individual ideas, when they occupy their own pitch area in the composite, textural density of the musical texture. This area can be large or quite small, and still be effective.

Sounds that do not have well-defined pitch quality, occupy a pitch area. These types of sound are noiselike, in that they cannot be perceived as being at a specific pitch. Such sounds may, however, have unique pitch characteristics.

Many sounds cannot be assigned a specific pitch, yet have a number of frequencies that dominate their spectrum. Cymbals and drums easily fall into this category. Cymbals are easily perceived as sounding higher- or lower-than one another. Yet a specific pitch cannot be assigned to the sound source.

We perceive these sounds as occupying a pitch area. We perceive a pitch-type quality in relation to the registral placement of the area in which the highest concentration of pitch information (at the highest amplitude level) is present in the sound, and in relation to the relative density (closeness of the spacing of pitch levels) of the pitch information (spectral components). We are able to identify the approximate area of pitches in which the concentration of spectral energy occurs, and are thus able to relate that area to other sounds.

Pitch areas are defined as the range spanned by the lowest and highest dominant frequencies around the area of the spectral activity. This range is called the bandwidth of the pitch area. Many sounds will have several pitch areas where concentrated amounts of spectral energy is occurring, with one range dominating and others less prominent. The size of the bandwidth and the density of spectral information (the number of frequencies within the bandwidth and the spacing of those frequencies) define the sound quality of the pitch area.

Dynamic Levels and Relationships

Dynamic levels and relationships have traditionally been used in musical contexts for expressive or dramatic purposes. Expressive changes in dynamic levels and the relationships of those changes have most often been used to support the motion of melodic lines, to enhance the sense of direction in harmonic motion, or to emphasize a particular musical idea. A change of dynamic level, in and of itself, can produce a dramatic musical event, and is a common musical occurrence. Changes in dynamic level can be gradual or sudden; subtle or extreme.

Dynamics have traditionally been described by analogy: louder than, softer than; very loud (*fortissimo*), soft (*piano*), medium loud (*mezzo-forte*), etc. The artistic element of dynamics in a piece of music is judged in relation to context. Dynamic levels are gauged in relation to

- (1) the overall, conceptual dynamic level of the piece of music,
- (2) the sounds occurring simultaneously with a sound source in question, and
- (3) the sounds that immediately follow and precede a particular sound source.

The components of dynamic levels and relationships in audio recording are dynamic contour (with gradual and abrupt changes in dynamic level), emphasis/deemphasis accents (abrupt changes in dynamic level), musical balance (gradual and abrupt changes in dynamic levels), and dynamic speech inflections.

Rapid, slight alterations or changes in dynamic level for expressive purposes are often present in live performances. This is called tremolo, and is used primarily to add interest and substance to a sustained sound. Tremolo and vibrato are often confused. Vibrato is a rapid, slight variation of the pitch of a sound;

it, also, is used to enhance the sound quality of the sound source. At times, performers may not be able to control their sound well enough to control tremolo and vibrato alterations; in these instances, tremolo and vibrato may detract from the source's sound quality, rather than contribute to it.

Changes in dynamic levels over time comprise dynamic contour. Dynamic contours can be perceived for individual sounds, individual sound sources, individual musical ideas comprised of a number of sound sources, and the overall piece of music. Dynamic contour can be perceived from many different perspectives: from the smallest changes within the spectral envelope through great changes in the overall dynamic level of a recording.

The composite of all of the dynamic contours creates musical balance. Musical balance is the interrelationships of the dynamic levels of each sound source, to one another and to the entire musical texture. The relative dynamic level of a particular sound source in relation to another sound source is a comparison of two parts of the musical balance.

Dynamic contours and musical balance have been used in supportive roles in most traditional music. At times dynamic level changes have been used for their own dramatic impact on the music, but most often they are used to assist the effectiveness of another artistic element.

To support a musical idea or to create a sense of drama, musical ideas are often brought to the listener's attention by dynamic emphasis or attenuation accents. A shift in dynamic level that brings the listener's attention to a musical idea, is an accent. Accents are most often emphasis accents, making use of increasing the dynamic level of the sound to achieve the desired result. Much more difficult to successfully achieve, deemphasis (or attenuation) accents draw the listener's attention to a musical idea, or a sound source, by a decrease in the dynamic level of the sound. Attenuation accents are often unsuccessful because the listener has a natural tendency to move attention away from softer sounds; these accents are most easily accomplished in sparse musical textures, where little else is going on to draw the listener's attention away from the material being accented.

Dynamic levels and relationships may be significantly different in the final recording than they were originally performed. The recording process has very precise control over the dynamic levels of a sound source in the musical balance of the final recording. An instrument may have an audible dynamic level in the musical balance of a recording that is very different from the dynamic level at which the instrument was originally performed. The timbre of the instrument will exhibit the dynamic levels at which it was performed (perceived performance intensity), but its relative dynamic level in relation to the other musical parts might be significantly altered by the mix. For example, an instrument may be recorded playing a passage at *ff*, with the passage ending up in the final musical balance at a very soft dynamic level; the timbre of the instrument will send the cue that the passage was performed very loudly, yet the actual dynamic level will be quite soft in relation to the overall musical texture, and to the other instruments of the texture.

The dynamic level of a sound source in relation to other sound sources, musical balance, is quite different and distinct from the perceived distance of one sound source to another. Yet, these two occurrences are often confused, and is the source of much common, misleading terminology used by recordists. Significant differences are present between a softly generated sound that is close to the listener, and a loudly performed sound that is at a great distance to the listener, even when the two sounds have precisely the same perceived loudness level. Loudness levels within the recording process are independently controllable from the loudness level at which the sound was performed, and are independently controllable from the distance of the sound source from the original receptor and from the person listening to the final recording.

Rhythmic Patterns and Rates of Activities

Durations of sounds (the length of time in which the sound exists) combine to create musical rhythm. Rhythm is based on the perception of a steadily recurring, underlying pulse. The pulse does not need to be strongly audible to be perceived. The underlying pulse (or metric grid) is easily recognized by humans as the strongest, common proportion of duration (note value) heard in the music.

The rate of the pulses of the metric grid is the tempo of a piece of music. Tempo is measured in metronomic markings (pulses per minute, abbreviated "M.M."), or in some contexts as pulses per quarter note. Tempo, in a larger sense, can be the rate of activity of any large or small aspect of the piece of music (or of some other aspect of audio, for example the "tempo of a dialogue").

Durations of sound are perceived proportionally in relation to the pulse of the metric grid. The human mind will organize durations into groups of durations, or rhythmic patterns. In the same ways that we perceive patterns of pitches, we perceive patterns of durations. Pattern perception is transferable to all of the components of all of the artistic elements, and is the traditional way in which we perceive pitch and rhythmic relationships.

Rhythmic patterns are the durations of or between occurrences of an artistic element. Rhythmic patterns might be created by the pulsing of a single percussion sound; in this way rhythmic patterns would be created by the durations between the occurrences of the starts of the same sound source. Rhythmic patterns comprised of the durations of successive, single pitches (perhaps including some silences) creates melody. Rhythmic patterns of the durations of successive chords (groups of pitches) creates harmonic rhythm. In this way, rhythm can be transferred to ALL artistic elements. For examples, it is possible to have rhythms of sound location (as is becoming a very common mixing technique for percussion sounds); it is likewise possible to have timbre melodies, or rhythms applied to patterns of identifiable timbres.

Sound Sources and Sound Quality

The selection, modification, or creation of sound sources is an important artistic element of audio recording. The sound quality of the sound sources (the timbre of the source), plays a central role in the presentation of musical ideas, and has become an increasingly significant resource for the expression of musical ideas.

The sound quality of a sound source may cause a musical part to stand out from others, or to blend into an ensemble; in and of itself, it can convey tension or repose, or lend direction to a musical idea; it can add dramatic or extra-musical meaning or significance to a musical idea; finally, the timbral quality of a sound source can, itself, be a primary musical idea, capable of conveying a meaningful musical message.

Until the Twentieth Century, composers of Western music used the sound quality of a sound source

- (1) to assist in delineating and differentiating musical ideas,
- (2) to enhance the expression of a musical idea by the careful selection of the appropriate musical instrument to perform a particular musical idea, or
- (3) to create a composite timbre (or texture) of the ensemble, thereby forming a characteristic, overall sound quality.

Performers have always utilized the characteristic timbres of their instruments or voices to enhance musical interpretation. This activity has been greatly refined by the resources of recording technology. The recording process allows the performers greater flexibility in shaping the timbre of their instruments for creative expression. Of equally great importance, after the performance has been captured, the recording process allows for the opportunity to return to the performance for further (perhaps extensive) modifications of sound quality.

The selection of a sound source to represent (present) a particular musical idea is vital to the successful communication of the idea. The act of selecting a sound source is among the most important decisions composers (and producers) make. The options for selecting sound sources are:

- (1) to choose a particular instrumentation,
- (2) modifying the sound quality of an existing instrument or performance, or
- (3) to create, or synthesize, a sound source to meet the specific need of the musical idea.

The selection of instrumentation was once merely a matter of deciding which generic instrument of those

available would perform a certain musical line. The selection of instrumentation has become very specific, since the performance of a recording may virtually live forever, whereas previous performance existed for only a passing moment.

Today, the selection of instrumentation is often so specific, as to be a selection of a particular performer playing a particular model of an instrument. Generally, composers and producers are very much aware of the sound quality they want for a particular musical idea. The performer, the way the performer can develop a musical idea through their own personal performance techniques, and their ability to use sound quality for musical expression are all considerations in the selection of instrumentation.

Vocalists are commonly sought for the sound quality of their voice, and their abilities to perform in particular singing styles. The vocal line of most songs is the focal point that carries the weight of musical expression. Vocalists make great use of performance techniques to enhance and develop their sound quality, as well as to support the drama and meaning of the text.

Performance techniques vary greatly between instruments, musical styles, performers, and functions of a musical idea. The most suitable performance techniques will be those which achieve the desired musical results, when the sound sources are finally combined. One performance technique consideration must be singled out for special attention: the intensity level of a performance.

A performance on a musical instrument will take place at a particular intensity level. This perceived performance intensity is comprised of loudness, performance technique and the expressive qualities of the performance. Each performance at a different intensity level results in a characteristic timbre of that instrument, at that loudness level. The same sound source will have different timbres, at different loudness levels (and at different pitch-levels), etc., through performance intensity.

Along with the timbre (sound quality) and the loudness level, comes a sense of drama and an artistically sensitive presentation of the music, that is communicated to the listener. Through performance intensity, louder sounds might be more urgent, more intense; softer sounds might be cause for relaxation of musical motion. Much dramatic impact can be created by sending conflicting loudness level and sound quality information; a loud whisper, a trumpet blast heard at pianissimo.

Modifying an existing sound source is a common way of creating a desired sound quality. Instruments, voices, or any other sound may be modified (while being recorded, or afterwards) to achieve a desired sound quality. Most often, this option for selecting a sound source is in the form of making detailed modifications to a recorded performance of a musical idea by a particular instrument. The final sound quality will still have the characteristic qualities of the original sound.

The extensive modification of an existing sound source, to the point where the characteristic qualities of the original sound are lost, is actually the creation of a sound source. The creation of new sound qualities (or inventing timbres) has become an important feature in many types of music. The recording process easily allows for the creation of new sound sources, with new sound qualities.

Sound qualities are created by either extensively modifying an existing sound (through sound sampling technologies) or by synthesizing a waveform. Sound synthesis techniques allow precise control over these two processes, and are having a widespread impact on recording practice and musical styles. Many specific techniques exist for synthesizing and sampling sounds; all with their own unique sound qualities and own unique ways of allowing the user to modify the sound source.

With the control of timbre by the recording process, has come a new sense of the importance of sound quality to communicate, as well as to enhance, the musical message. Sound quality has become a central element in a number of the primary decisions of recording of music, and in the creation of music through the recording process. In making these primary decisions, sound quality is conceptualized as an object. The sound is thought of as a complete and individual entity.

In this way sound quality is considered as a sound object. While the sound object is comprised of

component parts (as we have discussed above), it is perceived as a large unit, for its overall sound qualities.

Sound quality is perceived as a sound object:

- (1) when the sound quality of the sound source itself is at the center of the listener's attention, or
- (2) when the sound itself is the most important element of the musical texture.

For the sound object, the individual character of the sound source is significant. This is in contrast to the normal, primary significance of how the sound quality enhances the musical material, or how the sound sources relate to one another.

The entire sound of the music may also be conceptualized as a single entity, or overall quality. In this way, the overall musical sound is perceived as a large sound object, being comprised of any number of small, individualized sound objects.

This sound quality of the overall sound, or entire program is texture. As it is the composite of all sound objects present at any one time, over a span of time, texture has also been called sound structure or sound event.

Texture is perceived by the characteristics of its global sound quality. This concept of sound quality can be applied to groups of sound sources, in the same way as to individual sound objects or the entire program.

Texture will nearly always be comprised of any number or types of individual sound objects or groups of sound sources. Texture is perceived as an overall character that is comprised of the states and activities of its component parts. Pitch-register placements, rate of activities, dynamic contours, and spatial properties are primary factors in defining a texture by the states or values of its component parts.

The reproduced recording presents an illusion of a live performance. This performance will be perceived as having existed in reality, in a real physical space; as the human mind will conceive of any human activity in relationship to their own physical experiences. The recording will appear to be contained in a single, perceived physical environment. Within this perceived space is an area that comprises the sound stage.

The **sound stage** is the location within the perceived performance environment, where the sound sources appear to be sounding. The sound stage will appear to be contained within a single, global environment. The sound sources of the recording will be grouped by the mind, and will appear to occupy a more or less specific area of that global environment; this area is the sound stage. It is possible for different sound sources to occupy significantly different locations within the sound stage.

Imaging is the perceived lateral location and distance placement of the individual sound sources within the sound stage. Imaging is defined by the perceived physical relationships of the sound sources. As such, it is the perceived locations of the sound sources within the stereo array and with respect to perceived distance.

The stereo (lateral) location of a sound source is the perceived placement of the sound source in relation to the stereo array. Sound sources may be perceived at any lateral location within, or slightly beyond, the stereo array.

Phantom images are sound sources which are perceived to be sounding at locations where a physical sound source does not exist. Imaging relies on phantom imaging to create lateral localization cues for sound sources. Through the use of phantom images, sound sources may be perceived at any physical location within the stereo loudspeaker array, and up to 15° beyond the loudspeaker array. **Stage width** (sometimes called stereo spread) is the area that spans the boundaries established by extreme left and right images of the sound sources.

Phantom images not only provide the illusion of the location of a sound source, they also create the illusion of the physical size of the source. Two types of phantom images exist: the spread image and the point source.

The **point source phantom image** is a focused, precise point in the sound stage where a sound source appears to be located. It is an exact location where the sound source is perceived as occupying. It appears to have a physical size that is quite narrow, and precisely located in the sound stage.

The **spread image** appears to occupy an area. It is a phantom image that has a size that extends between two audible boundaries. The size of the spread image can be considerable; it might be slightly wider than a point source, or it may occupy the entire stereo array. The spread image is defined by its boundaries; it appears to occupy an area between two points. At times, a spread image may appear to have a "hole in the middle," where it might occupy two equal areas, one on either side on the stereo array.

The perceived lateral location of sound sources can be altered to provide the illusion of moving sources. Moving sound sources may be either point sources or spread images. Point sources that change location most closely resemble our experiences of moving sound sources.

The listener will perceive two types of distance cues from the recorded music:

- (1) the distance of the listener to the sound stage, and
- (2) the distance of each sound source from the listener. Both of these distances rely on a perception that the entire recording occupies a single, global environment. This perceived performance environment establishes a reference location of the listener, from which all judgments of distance can be calculated.

Spatial Properties

The spatial properties of sound have traditionally not been used in musical contexts. The only exceptions are the location effects of antiphonal ensembles of certain Renaissance and early Twentieth-Century musics, and the effect of the movement of the sound source found in certain drama-related works of the Nineteenth Century.

The spatial properties of sound play an important role in communicating the artistic message of recorded music. The roles of spatial properties of sound are many: it may be to enhance the effectiveness of a large or small musical idea; it may help to differentiate one sound source from another; it may be used for dramatic impact; it may be used to alter reality or to reinforce reality.

The number and types of roles that spatial location may play in the communication of a musical idea have yet to be exhausted or defined.

All of the components of the spatial properties are under very precise and independent control. All of the spatial properties have the capacity to be in, and to gradually change between, many dramatically different and fully audible states.

The spatial properties of sound that are of primary concern to recorded music (sound) are:

- (1) the perceived stereo location of the sound source on the horizontal plane of the stereo array,
- (2) the conceptualized distance of the sound source from the listener, and
- (3) the perceived characteristics of the sound source's physical environment. The perceived elevation of a sound source is not consistently reproducible in widely used playback systems, and has not yet become a resource for artistic expression.

The three spatial properties are realized through stereophonic sound reproduction. The spatial attributes are related by the perceived relationships of location and distance cues of the sound sources in relation to the sound stage, and the relationships of the sound stage to the perceived performance environment of the recording.

Two-channel sound reproduction has become the standard for the recording industry, with monophonic

capabilities still considered for adoption for AM broadcast and television sound. The two-channel array of stereo sound attempts to reproduce all spatial cues through two separate sound locations (loudspeakers), each with more-or-less independent content (channel). With the two channels, it is possible to create the illusion of sound location at a loudspeaker, in between the two loudspeakers, or slightly outside the boundaries of the loudspeaker array; location is limited to the area slightly beyond that covered by the stereo array, and to the horizontal plane. The characteristics of the sound source's environment and distance from the listener are affected in much more subtle ways by the stereo reproduction format.

A setting is created by the two-channel playback format for the recreation of a recorded or created performance (complete with spatial cues). The setting of the two-channel playback format is a conceptual (and physical) environment within which the recording will be reproduced more-or-less accurately.

The stage-to-listener distance establishes the location of the sound stage with respect to the listener. It is the distance between the grouped sources that comprise the sound stage and the audience. This stage-to-listener distance is the placement of the sound stage within the overall environment of the recording, in relation to the perceived location of the listener

The depth of sound stage is an area occupied by the distance of each sound source relative to one another. The boundaries of the depth of the sound stage are the perceived nearest and the perceived furthest sound sources. The perceived distances of sound sources within the sound stage may be extreme.

The two factors of distance cues interact. The depth of the sound stage is perceived in the context of the stage-to-listener distance; the listener is prone to placing the nearest sound source at the nearest location of the stage-to-listener distance. Conversely, the perceived distance of each sound source relative to the listener can cause a shift in the perceived stage-to-listener distance; especially in multi-track recordings that incorporate dramatic reverberation techniques.

These two factors of distance cues have different levels of importance in different contexts. Depth of sound stage cues tend to be emphasized over stage-to-listener distance cues in many recordings; the cues of the distance of the source from the listener are often exploited to support dramatic and musical ideas. As another example, stage-to-listener distance cues are carefully calculated in many art music recordings (especially those utilizing standardized stereo microphone techniques); while the distance might not change within the recording, the stage-to-listener distance is carefully selected to represent the most appropriate vantage point (the ideal seat) from which the music is to be heard.

The matching of a sound source to the sound characteristics of an environment in which it will sound, and the selection of the environment of the sound stage (the perceived performance environment) have become important parts of music recording. This coupling of source to environmental characteristics has the potential to have a significant impact on the meaning of the music, of the text, or of the sound source; to supply dramatic effect; to segregate sound sources, musical ideas, or groups of instruments; to enhance the function and effectiveness of a musical idea.

The sound characteristics of the host environments of sound sources and the complete sound stage are precisely controllable. Each sound source has the potential to be assigned environmental characteristics that are different from the other sound sources. The recording process allows for the assigning of different environments to different sound sources, and for widely varying those characteristics as desired. Further, each source may occupy any distance from the listener within the applied host environment.

The environment of the sound stage and individual environments for each sound source (or groups of sound sources) often co-exist in the same music recording. This musical context places the individual sound sources with their individual environments "within" the overall environment of the recording. The result is a perception:

- (1) that physical spaces may exist side-by-side,
- (2) that one physical space may exist within another physical space (to the point where a larger physical space may be perceived to exist within a smaller physical space), and

(3) that sounds may exist at various distances within the same or different host environments, within other environments (causing conflicting distance cues between sources and environments). The result is the illusion of space within space.

Any number of environments and associated stage-depth distance cues may occur simultaneously, and coexist within the same sound stage. The environments and associated distances are conceptually bound by the spatial impression of the perceived performance environment. These "outer walls" of the overall program establish a reference (subconsciously, if not aurally) for the comparison of the sound sources. Oddly, the overall space that serves as a reference, and that is perceived by the listener as being the space within which the other activities occur, might have the sound characteristics of a physical environment significantly smaller than the spaces of the sound sources it appears to contain. Such cues that send conflicting messages between our life experiences and the perceived musical occurrence can be used to great artistic advantage. This is a very common space within space relationship.

Space within space will at times be coupled with distance cues to accentuate the different environments (spaces) of the sound sources. Often, this illusion will be created solely by the environmental characteristics of the different spaces of each sound source.

Space within space has become an important element in shaping the imaging of a recording. Often, imaging will work in a complementary and contrasting fashion with musical balance. Recordings are often quite sophisticated in the interaction of these two artistic elements.

With the recording process, it is possible for any of the artistic elements of sound to be varied in considerable detail. In so doing, all artistic elements can be shaped for artistic purposes, or used as resources for artistic expression. As all elements of sound are capable of an equal amount of variation, it is possible for each element of sound to function in any role in communicating the message of a piece of music.

The artistic elements are used in very traditional roles in certain musical works and types of recording productions, and in very new ways in other works. The new ways the artistic elements are used tends to place more emphasis on the artistic elements of sound that can not be precisely controlled acoustically. Many new musical compositions use the artistic elements unique to audio recording (especially sound quality and spatial properties) to support the musical material. Different musical relationships and sound properties exist in audio recordings than can found in the music conceived before the artistic resources of recording technology existed.

<http://www.diyaudio.com/forums/multi-way/183337-what-ment-term-soundstage.html>

<http://forums.audioreview.com/showthread.php?t=9713>

<http://www.stereophile.com/reference/50/index.html>

http://philomel.com/pdf/PsychMus_Ch9.pdf

http://philomel.com/phantom_words/

Recommended reading:

http://deutsch.ucsd.edu/psychology/deutsch_publications.php

