Localization Cue Correction (LCC): The Rationale

A technology developed at
L&L Innovation Labs (www.L2iL.com)

I

Alan Dower Blumlein, an EMI sound engineer, filed a British Patent Specification 394325 (around 88 years ago) entitled "*Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing systems*." Blumlein told us that his reproduction method using two widely spaced loudspeakers was flawed, yet we still use the ~88-year-old *stereo triangle* method he deemed unsatisfactory. A reconsideration of this psychoacoustic issue holds great promise for the move into more virtualized content experiences.

Blumlein stated that the maximum width of the stereophonic sound image is narrowed to the angle the speakers subtend at the listening point. The sense of direction of the ostensible sound source will only be transferred to a listener for the full frequency range for locations lying between the loudspeakers. Hence, it is not a truly binaural event. Thus, the last major hurdle in having truly realistic sound reproduction is the continued use of the ~88-year-old 60-degree loudspeaker triangle and its home theater cousins. The use of speakers in this position (and similar wide angles) results in three varieties of localization cue distortion. The first of these concerns the outer ears or pinnae. These cavities and ridges produce peaks and valleys in the frequency response in the region above about 1500 Hz. The brain uses these patterns to determine the point of origin of a sound, just like radio direction finders. These patterns are individual like fingerprints; ergo, any quality reproduction system must include elements that do not interfere with this localization mechanism or try to synthesize them using filters. If a speaker is at 30-degrees, then all the pinna cues are always 30-degrees, and this contradicts other localization cues— and is also abnormal. Thus, the first step in fixing stereo or surround systems is to move the frontal (and rear) speakers closer together[1] to an angle between 5-25-degrees.

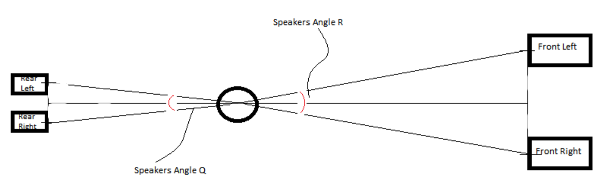


Figure 1 (4.0 isosceles triangle use case)

---

[1] Instead of the common stereo-triangle (equilateral triangle setup), a perfect stereo setup (using LCC) will look more like an isosceles triangle, as seen in figure 1.

This guarantees that the pinna function is more natural for center instruments, SFX, or dialog in movies, and obviates the need for a center speaker[2] in home theater systems— and in many cases, rear channel speakers. Note, that when sounds come from the side, they have a direct path to the ear canal; therefore, the pinna direction finding mechanism is less critical at wider angles (Optionally, the use of two rear speakers can correct for this).

Moreover, having the speakers closer together eliminates the head shadow, i.e. a portion of a listener's HRTF, wrongly introduced when speakers are at a wider angle. In normal hearing, the effect of the head shape on interaural levels and delays change with the position of a sound; however, in stereo, there is a fixed shadow of 30-degrees from each side no matter where a recording says an instrument is supposed to be. So, by moving the speakers more or less in front of the listener, the listeners head shadow now has nearly no effect on localization cues even with head rotations, nodding, or leaning (note: toeing speakers in when using LCC usually adds no real significant benefit. We tend to face speakers straight out– no toeing-in).

Though, there would be no stage width with speakers so close together unless speakers behaved like laser beams— one beam aimed at each ear, but that would drastically reduce the number of listeners that could enjoy some wider stage effect and even constrain one listener's ability to move their head around freely while still enjoying proper sound imaging. Localization Cue Correction, or LCC, is a *Recurrent Correction Sample Exchange Protocol*, that can be installed in speaker connected devices such as: laptop/desktop computers, tablets, OTT/Set-top boxes, Smart TV sets, microchips, automobiles, micro/game consoles, over radio frequency transmissions, or VR/AR systems to ensure that an ordinary pair of frontal and even with additional rear speakers do indeed behave like acoustical laser beams without the negative off-axis artifacts of directional loudspeakers. No head tracking, head response measurements, or other processing applications are needed. Indeed, the frequency response of an LCC equipped system is much flatter than the 60-degree stereo equivalent, which additionally, in many cases, reduces the need for digital loudspeaker/room correction and physical acoustic material treatments when front and rears are used together in a 2.0 extension or 4.0 surround setup. At lower frequencies, where the pinnae do not operate, the brain localizes sound by interaural level and time differences (ILD/ITD) between the ears. With speakers at 60-degrees, the maximum time difference between the average ears is about 220µs[3]. Compare this with about 750µs for a sound at the far stage side heard in, for example, a concert hall or picked up by a two-channel microphone. Therefore, if a recording actually recorded an instrument at the far side of the stage so the time difference is 750µs on the content file, that value will be truncated to an ITD of ~220µs when reproduced the traditional way; however, content will be maintained at 750µs[4] if LCC is used to process the 2.0 (or

[2] Center channel speakers are not recommended with LCC, as they, specifically with 2.0 content, narrow the sound stage. In any case, they are simply not needed with an LCC enabled system.

[3] Widely spaced standard stereo playback has a maximum ITD of 220µs (never wider than 60º), which is why many people are instructed to place their speakers at 60º angles. That means, the speaker to speaker distance is the same distance as the head to speaker distance (just imagine the shape of an equilateral triangle, sometimes known as the *stereo triangle*)

[4] The ITD should of course be ~700µs, and it can never be more in normal hearing since that is the maximum possible delay between the ears. The signal at the left ear should

multi-channel) source file over the arranged speaker placement stated earlier.

Speakers at 60-degrees also distort the recorded inter-channel level differences in a similar way, thus reducing the level differences captured by microphones again by ~2/3. This is why you can never get a stage wider than 60-degrees when using the stereo 60-degree reproduction equilateral stereo triangle. What also happens is that the brain is confronted with pinna, level, and delay cues that are inconsistent, which simply do not exist in nature. This is why listeners can so easily detect that they are listening to a recording among many other factors. LCC delivers to the ears whatever TD and LD were on a recording, and these cues are now largely consistent with the pinna direction finding mechanism.

Using LCC, you can expect to experience sound stages of up to 180-degrees wide for many recordings with enhanced depth and clarity. If you use a similar LCC processed second pair of speakers at the rear (at a 5-25-degree angle) you can have a full circle of direct sound using just four speakers. As mentioned, this rear LCC pair can also produce a sense of envelopment when used to play 2.0 files.

Note, that since the speakers in the stated set-up are relatively close together (less than ~90cm apart), the optimum listening area is smaller in width; but, it is much longer in length compared to stereo. Ergo, LCC allows for a better off-axis listening experience (LCC's offside listening provides excellent mono; so again, a center speaker in 5.1 is never needed).

II

Most attempts to render psychoacoustically accurate VSS (Virtual Surround Sound) or 3D audio over loudspeakers do not work for a variety of reasons. Efforts that have tried to achieve what LCC does or other VSS techniques do not generate true 3D experiences with a flat frequency response at the pinnae. For example, many attempts are not effective over multi-speaker soundbars and work mostly for content (mostly movies) that is encoded to reproduce a "3D effect". That is, attempts to achieve what LCC can do really only work in the slightest and for encoded content.

Most importantly, all VSS attempts miss pinnae adjustment possibilities. And, in some attempted VSS technologies, there is only one ILD and ITD reflection value from each speaker no matter what or where sound is playing. Many VSS efforts do not have a fully amendable method (such as our proprietary technology, e.g. LCC's *Computationally Recurrent Correction Sample Exchange Protocol*). Attempted VSS technologies, more often than not, require a fixed angle of the speakers. Many VSS attempts use HRTFs, which are redundant and thus unnecessary forms of DSP to render over loudspeakers, and/or head tracking to make things "work". Additionally, these attempts do not work over full surround.

In concert halls, you are lucky to have 5dB of interaural level differences. LCC can easily achieve 10dB, so it is quite adequate for any home media you can find; nonetheless, this is not the only spec that matters.

Most VSS attempts require fixed speaker angles, miss pinnae correction possibilities, and generate a poor frequency response due to combing. Many forms of VSS DSP make a fundamental mistake in confusing what is going into a speaker with what is heard at the ear, at the listening position. An LCC enabled outputs say the left signal followed by an enumerated amount of delayed, and attenuated, and polarity reversed samples of the input signals. All these samples are flat since they are not

reach the right ear in ~750µs as in normal hearing (or with LCC) rather than in 220µs as in the stereo triangle.

filtered in any way. However, if you use an averaging RMS frequency response meter to measure this train, you will not see a flat response but a response that keeps changing with time and the audio content. This is like trying to measure the frequency response at a seat in a concert hall. The response will be far from flat and will vary with where the sound from the stage originates and with what, for example, an instrument is playing. Measuring ordinary 60-degree loudspeaker stereo signals at the ear rather than at a speaker will show the same poor frequency response. That is one of the things LCC is designed to fix.

Now, these delayed samples that are reaching the ear to correct for cue localization mix everywhere in the room once you get a foot or so from the speakers; therefore, the response, even away from the listening area, is flat, and you hear both channels clearly, even if far off center. Ergo, even room reflections will not be colored, since the reflections mix all this together with more delays and attenuations[5]. What it really amounts to is that the direct signals are accompanied by a bunch of rich early reflections just like in a real acoustic space. If you use rear speakers, to accommodate surround sound or achieve immersive stereo, you can actually control this effect directly and produce full envelopment without drastic dependencies on room acoustic treatments and digital room/loudspeaker correction/calibration. In other words, the walls seem to vanish more or less. Moreover, LCC samples, after the direct sound, are a long sequence of early reflections that get launched repetitively from the front (and optionally rear) speakers and go all around the room lasting a very long time since LCC keeps providing samples at naturally arranged declining levels and increasing intervals, which is reminiscent of a real acoustic space. Eventually, the field in the small room is diffuse as in real acoustic spaces. Thus, you are surrounded by natural decaying reflections for tens of milliseconds per sample. The direct sound that reaches the ears first is also being reflected around the room, however, now it is followed by so much more deferred energy that it is harder for the brain to judge the room size the listener is in. Since most speakers are not very directional, there is even more LCC activity with different delays and directional interest persisting far longer than the room response time (RT of .2 sec for a typical small room). In non-LCC stereo playback systems, only one sound sample stimulates the room. This may explain why omnidirectional speakers work well since they do provide a slightly richer reflective room field than smaller highly directional speakers [6], which generate less reflective variety. Recorded ambience normally does not seem to help much since it is static and low level in 2.0 media. LCC processing enriches the reflective field and thus the sense of not being in a small room and/or a room needing much calibration.

Given our software algorithm solves the most pressing psychoacoustic and speaker physics issues in a purely scientific fashion and with elegant precision that no other VSS attempt has been able to achieve, audio/sound engineers (pro audio users) that produce content for music, movies, and/or games can, in addition to end consumers of audio content, benefit greatly from LCC. For example, translatability and insight into audio content (the mix/master) is of great

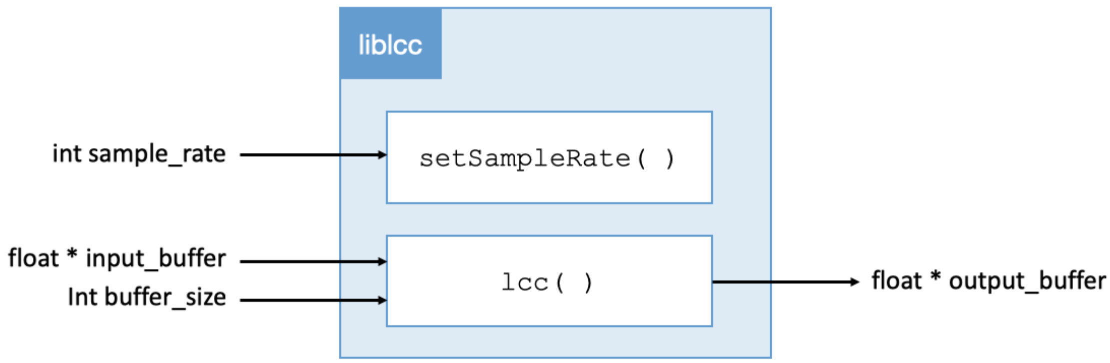
[5] Now these delayed samples mix everywhere in the room once you get a foot or so from the speakers, so the response even away from the listening area is flat and you hear both channels clearly, even if far off center. So, even room reflections will not be colored since these reflections are a mixture of all these samples with even more delays and attenuations, thus averaging out to be flat just as in a concert hall.

[6] Electrostatic/planar magnetic panels (vertical line arrays) fair well too.

importance to and is one of the things that haunts many musicians, producers, engineers, etc. in getting the final product ready for mass consumption over all speaker types. Many pro audio users find ways to help themselves overcome the hurdles of translatability in their mixes/masters in various ways. A few ways they can benefit from LCC, which no other means currently can provide, is by using an LCC enabled speaker system**.** LCC makes it **1)** easier to hear accidental reversed polarity of spot mics, cables, etc. (and what your panning algorithm is doing). **2)** There is near zero comb filtering, so you hear the true frequency response of what you are doing in the mix, i.e., LCC enabled systems reduce coloration by eliminating the peaks and dips in the higher frequency ranges (specifically above 2 kHz) and substantially reducing the low bass volume doubling induced when both stereo speakers (or 5.1+) send the same bass signal to both ears. **3)** Unlike stereo loudspeakers which have poor L/R channel separation, headphones/earphones provide complete isolation of the L/R channels (in fact, they have too much L/R isolation, which is one reason these devices have in-head localization issues). Thus, a pro audio engineer who uses speakers to decide on a stereo mix may not produce a product which sounds ideal using headphones/earphones. An LCC equipped loudspeaker closely mimics the separation achieved using headphones/earphones, thus allowing audio content producers to tailor their mixes/masters to the headphone and loudspeaker market. The advantage of using an LCC speaker enabled system for such advanced mixing applications is that the sound is externalized providing more natural auralization, and the pinnae are not unpredictably affected by the headphones/earphones structure. **4)** When not listening with your own head shadow but with the ILD/ITD head shadow created by the recording/mics used to produce the content, it makes it less likely that the final stereo pair mix will be distorted by the unavoidable features of your own head shadow. LCC systems used at pre/post audio production to end consumer electronic product stages presents a first-ever “end-to-end sonic stack” that brings about symmetry and immersion across loudspeaker playback systems. **5)** LCC reduces center bass doubling problems below ~90 Hz (In stereo, this mono bass is doubled since both speakers reach both ears exactly in phase and at the same level. Thus, the energy at each ear is double what it is when only one speaker has bass energy).

Additionally, some attempted VSS efforts apply frequency response filters so that the RMS response at the speaker terminals is the response at the ears and correct this. Of course, if one attempts to flatten at the speaker terminals, then it cannot be flat at an ear or do all the cue correction properly that LCC achieves. Thus, to correct for this, some VSS efforts measure the response at the ears and corrects this error. Of course, now the speaker signal is not flat, but that is as we now both know is inaudible. However, when you measure things at an ear you often end up with an unwanted pinna response flattening. That is a problem the LCC algorithm also avoids.

Easy-to-integrate C/C++ library




```
void setSampleRate(int sample_rate);

void lcc(float * output_buffer, float * input_buffer, int buffer_size);
```

- Partnering company’s engineers simply deploy an input and output buffer and wrap it around our API (written in C). This is about 1 hours’ worth of work for one engineer including E2E testing
- LCC must be placed last in the audio signal flow chain, e.g. after any audio DSP (ideally, no other spatial DSP in audio chain)
- LCC must not be routed to headphone jack or not be activated when headphones are used with system housing LCC algorithm
- LCC should be rerouted when user records audio to file, i.e., LCC should not be applied to recorded file (only used in real-time)
- Assuming engineering has a clear understanding of their audio buffers, 2-3 lines of code is pretty much all that is needed to connect to the LCC library
- LCC_app usage specs: RAM: ~3.4 MB and CPU: ~0.7% CPU (tested on a 2.7 GHz Intel Core i7). Chip and/or firmware integration would use much less power
- The buffer size determines the memory footprint. We can use a small buffer size with the algorithm running in real-time. LCC simply needs high-priority

Note on post-processing v audio encoding

Audio file encoding creates hard limits to the upper and lower bounds of the data representation (See [1]). As LCC manipulates the audio signal, values may occasionally (and briefly) be outside of the nominal range. To reduce clipping, audio signals must be scaled down before encoding to a file. This reduces dynamic range.

Typical audio systems allow ~+3dB headroom for post-processing effects. For example, Android's documentation specifies that smartphones/tablet implementations will provision for this +3dB headroom (See [2]). LCC fits comfortably within the +3dB range. This allows LCC to preserve the original dynamic range of the audio file while creating the LCC effect for post-processing audio systems.

[1] https://developer.android.com/reference/android/media/AudioFormat
"The nominal range of ENCODING_PCM_FLOAT audio data is [-1.0, 1.0]. ... Values outside of the nominal range are clamped before sending to the endpoint device."

[2] https://developer.android.com/reference/android/media/AudioTrack
"The implementation does not clip for sample values within the nominal range [-1.0f, 1.0f], provided that all gains in the audio pipeline are less than or equal to unity (1.0f), and in the absence of post-processing effects that could add energy, such as reverb. For the convenience of applications that compute samples using filters with non-unity gain, sample values +3 dB beyond the nominal range are permitted."

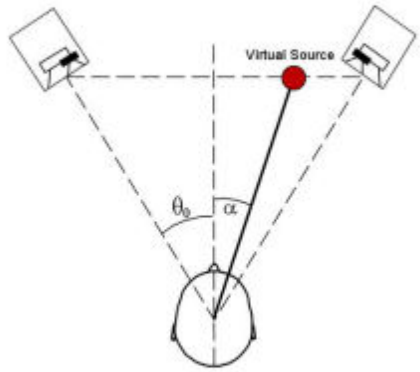
The philosophy:

The lowest common denominator for all sound playback systems is the signal that is feed to it. It is an incorrect view that audio systems and their sound radiator components need to "create" the spatial impression with the aid of its speaker positions or by HRTF/VSS-based effects or simply through reflecting sound by relying on the room boundaries to generate an immersive experience no matter what the genuine recording sounds like. The philosophy behind LCC’s science is to deliver a sound experience that is as true to the inter-channel relations of the sound recording as possible.

There is more to stereo than left and right (and phantom center) channels. The word stereo comes from the Greek word meaning solid, which implies the three-dimensions— height, width, and depth— in other words, spatial sound.

The most frequent spatial data is conveyed via two channels (ordinary stereo).
This is defined by a left and a right electrical signal that are partly same and partly different.
The equal section of the electrical audio signal describes the depth of the recording.
The variations describe the width of the recording.
The equal section and the differences describe a common floor area or zone.

Traditional thinking of spatial audio playback consists of two or more distinct loudspeaker drivers that are localized by the listener.
The spatial audio signal is then reproduced on *top* of these pre-determined positions of the separate speakers.
The most recognized theory around this expresses the math, valid up to 700Hz, to be as follows[7]:



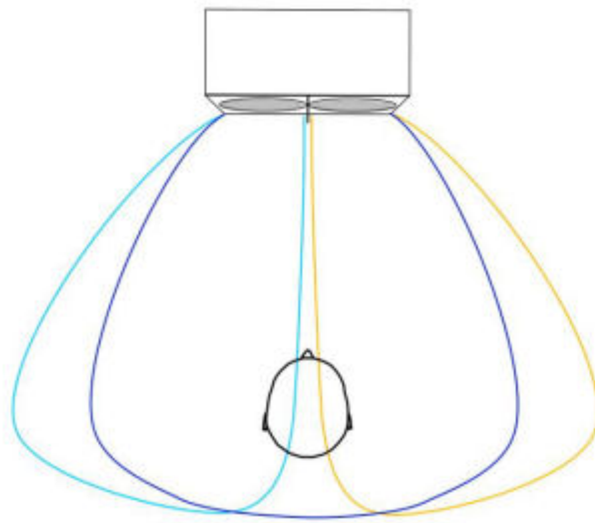
$$[(L-R) / (L+R)] \sin \theta_0 = \sin \alpha$$

where:
α is the angle of the percieved virtual image,
L-R is the directional information,
L+R is the mono information,
θ_0 is half the angle separating the loudspeakers.

With this context, it is clear that the directional data in the recording is in no way true to the perceived angle of the virtual sound source except for when a left only and/or right only signal is to be reproduced. All other directional data in the audio recording is smeared by the physical sound radiator placement and dispersion/directivity patterns. The more dominant the left and the right signal gets, the worse the smearing. This results in a sound image that produces an extensive mono effect.

[7] Ref: John William Strutt, 3rd Baron Rayleigh's Duplex theory to learn more.

An LCC optimized speaker system, as seen in the figure below, ideally utilizes two (or four for surround sound applications) mid-range + tweeter driver array units placed in a cabinet in front of the listener (and one behind the listener for surround sound content). Any number of subwoofers desired is not applicable (user preference). This makes it viable to reproduce all spatial data, forgoing any pre-determined positions of separate speakers, as seen in the figure below.



The aim of LCC is to reconstruct the spatial information from the recording as faithfully and truthfully representative of the audio signal as possible. The proprietary LCC algorithm rebuilds and purifies the inter-channel relationships reproduced by means of the speaker playback system around it in a way that is tailored properly to the biological sound detection finders (pinnae). As different from common virtualization technologies, LCC reproduces a multi-dimensional soundstage. To make this happen, the audio signal must be tailored to the speaker configuration. You will experience precisely what is in the recording...just as it was in the recorded material with width, depth, and height intact.

It is common knowledge that the typical consumer rarely ever places their left and right speakers correctly and that this problem is only compounded by multi-speaker setups, such as surround sound systems. LCC eliminates these variables and works best when housed in a standard reference monitoring system that can be relied upon without the chance of incorrect placement.

An LCC enabled speaker playback system will render a sharper, clearer, higher-definition sound imaging while, at the same time, eliminating the fatigue from processing redundant L/R signal information.

LCC is the definitive spatial audio solution. It is a timeless approach to spatial audio for speaker connected devices. As long as we use speakers to playback audio, LCC will be the most psychoacoustically relevant technology and carry any product that benefits from it into the future.

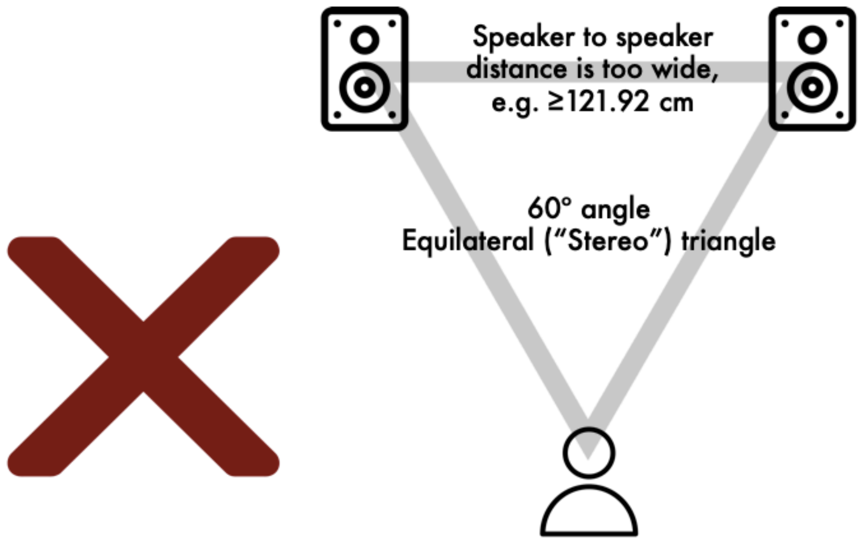
High-level product specification and design considerations:

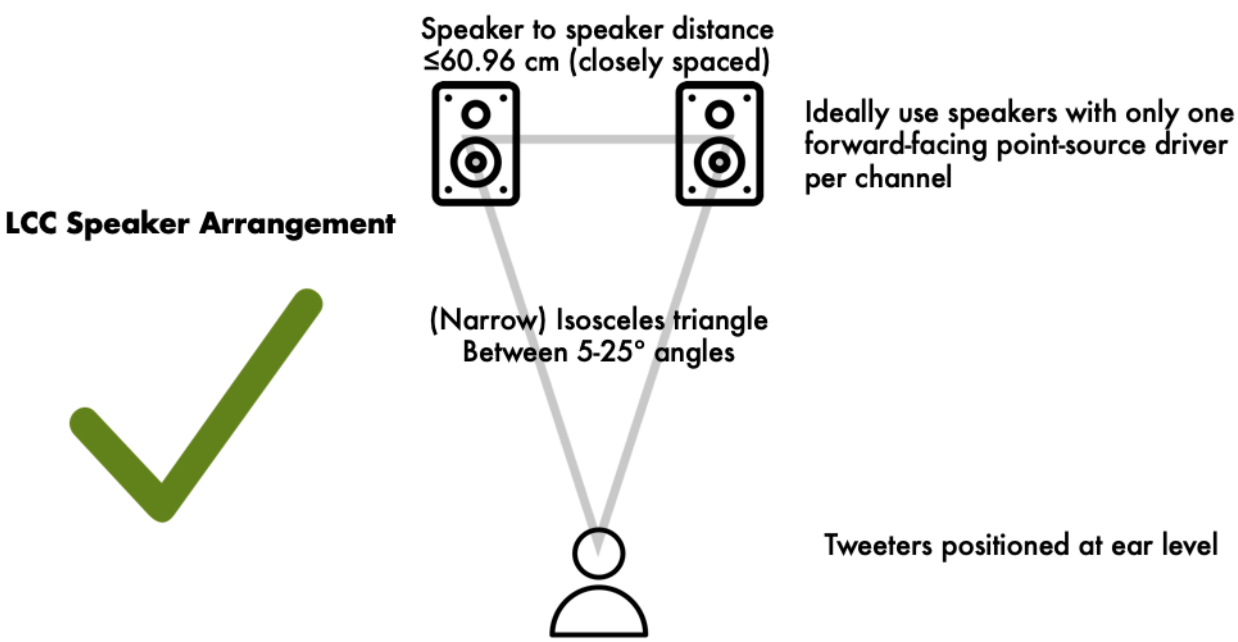
- Stereo imaging remains intact off-axis; thus, extending the optimal listening area
- Produces a strong phantom center (so, no need for a center channel)
- The perceived width of the stereo image increases proportionally to the listening distance
- Rear channel speaker distance can be digitally positioned with an optically integrated camera to measure the millisecond delay parameter, which allows user's rear channels to be digitally placed along various perceived physical distances to accommodate for possible room dimension limitations
- Uniquely separated and treated (transmission line) volumes in one cabinet
- Better translatability across various speaker systems and headphones:
    - Keeps sum and difference information intact better than vectors Left and Right, enabling the listener to experience all information originally captured in the recording. Also, LCC reduces listener's head shadow effects, which fosters deeper insights into the ILD/ITD head shadow created by the recording/mics used to produce the content
    - LCC reproduces a ~750µs ITD, unlike frontal standard stereo limited at 220µs ITD, optimizing for what a human head can produce (~700µs ITD)
    - Near zero comb filtering delivers an extremely flat frequency response
    - Reduction in the center channel bass doubling effect
    - Self-contained room/loudspeaker correction system (generated after LCC in to avoid resonance overcompensation (LCC placed before the power amp stage))
    - Reproduces all spatial cues and information that were encoded into the audio content, bringing to life a realistic 180˚ continuous and natural binaural soundstage of frontal 3D/immersive audio
    - Easier to hear accidental reversed polarity of spot mics, cables, and what the panning algorithm is/was doing

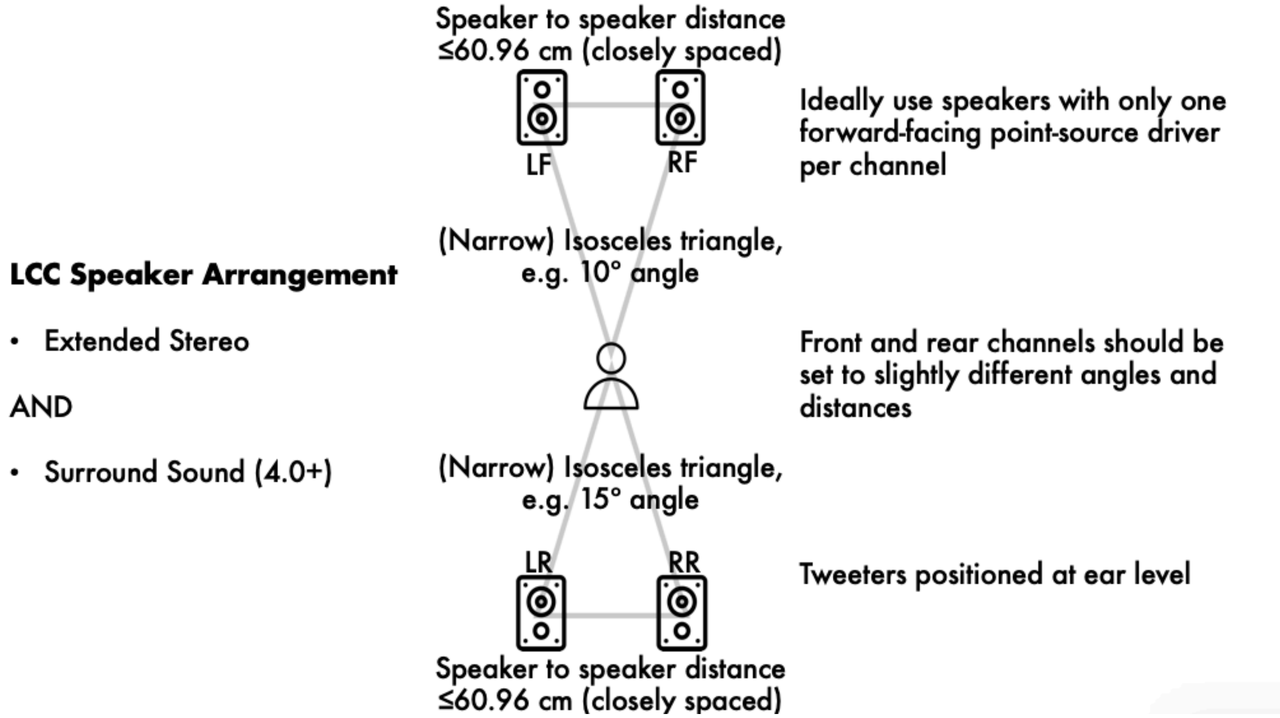
Additional specification and design considerations:

- Active full-range drivers, i.e., point source drivers for precise time alignment with matched directivity in all directions (see KEF's Uni-Q® technology or Genelec's Minimum Diffraction Coaxial (MDC™) Driver Technology as examples)
- Transmission line MDF enclosure (to stop overshoot and ringing and allow drivers/woofers to move back to their neutral points (see PMC's ATL Technology as an example) with an integrated 9" dual force-canceling woofer. Arranged as a satellite speaker system to save cabinet space (two separate units: 1. Unit with mid-range and tweeter and 2. unit with woofer/sub)
- The possible smallest size of playback speaker system might be measured as large two KEF LSX's stacked next to one another (not including the separate woofer unit)

Driver Placement Figurations






LCC Speaker Arrangement
• Extended Stereo
AND
• Surround Sound (4.0+)
Speaker to speaker distance ≤60.96 cm (closely spaced)
LF
RF
(Narrow) Isosceles triangle, e.g. 10° angle
(Narrow) Isosceles triangle, e.g. 15° angle
LR
RR
Speaker to speaker distance ≤60.96 cm (closely spaced)
Ideally use speakers with only one forward-facing point-source driver per channel
Front and rear channels should be set to slightly different angles and distances
Tweeters positioned at ear level