BY MIKE WOOD

Sound and light

ONE OF THE BOTH JOYS AND WOES of writing this column is that I have no rules. The editor of *Protocol* doesn't tell me what to write about; in fact she doesn't usually know my current topic until she sees copy! That should make life very easy for me, just think of a topic, then write about it. Therein lies the problem, just think of a topic . . ..

Oftentimes, topics arise organically as a result of a question somebody asks me, or as part of an ongoing theme such as the "How do LEDs work" series. I have a list I keep of possible topics, and there's quite a lot on that list, but none of them inspired me when I sat down in front of that blank page this month.

> **". . . you couldn't design a poorer system for discriminating wavelengths."**

As I was staring blankly at the screen, and listening to some music, I started daydreaming about what the world would be like if our vision system were like our hearing and vice-versa. Okay, that's potentially interesting. Let's make that the topic. I have no idea where this is heading, and no momentous goal in mind, so forgive me as I ramble. I'm making this up as I go! I'm sure this topic has been discussed somewhere before, but I couldn't find any reference to it in my research, so what you read here, along with mistakes and warts, is all mine. Don't trust a word of it!

Our hearing and vision systems are both critical to the way we perceive the world, and both evolved over the millennia in response to our activities in response to our environment. Helpful traits propagate, unhelpful ones die out. All very fine, but those two sensory systems are completely and utterly different.

I don't mean just the obvious differences that one is a system that responds to light while the other responds to sound. Both of those are waveforms of one kind or another, so you might imagine we might have similar mechanisms and processing, even though the actual wavelengths are so different. However, we don't; we have completely different discriminatory systems that are wired and processed in very different ways.

Let's start with hearing, as that seems to be the older sense. The human ear, although a poor example compared to other animals, has a quite incredible range. It can recognize and distinguish more than 12 orders of magnitude of dynamic range in sound sensitivity, and three orders of magnitude in frequency response. The eye also has a fantastic dynamic range, not quite as good as the ear, but still very impressive at about six orders of magnitude (That's over 20 stops if you are used to cameras). In fact, if it happens to hit a receptor head on, we have the ability to detect a single photon with our eyes. Can't get any better than that.

Where the eye differs hugely from the ear though is in the frequency response and the discrimination accuracy within that frequency range. As I mentioned before, the ear (at least for someone much younger than I) has a frequency range of about three orders of magnitude, roughly 20 Hz to 20 kHz. Our vision system has a much narrower range. We can see wavelengths roughly ranging from 390 to 700 nm, which is equivalent to a frequency range of 430 THz – 770 THz (THz is Terahertz - or 10^{12} Hz). This is only about a 2:1 range, much less than even one order of magnitude.

> **"I can't show you a picture of all this, you have to treat it as a thought experiment and imagine it."**

The reason for this limitation in vision frequency response is perhaps obvious, our eyes developed under the light from the sun and it makes no immediate sense to be able to see anything outside that band. **Figure 1** shows the normal electromagnetic spectrum with the range of human vision indicated. If we had the same three orders of magnitude range in vision that we do in hearing we might be able to see from 20 nm – 20 µm. That's extreme ultraviolet through to mid-infrared. However, most of that doesn't make it through our atmosphere so it would be black. This range is nowhere near far enough to get into potentially interesting microwaves or radio frequencies. So, on the face of it, it looks like there wouldn't be much point in having our vision extend much further than it does.

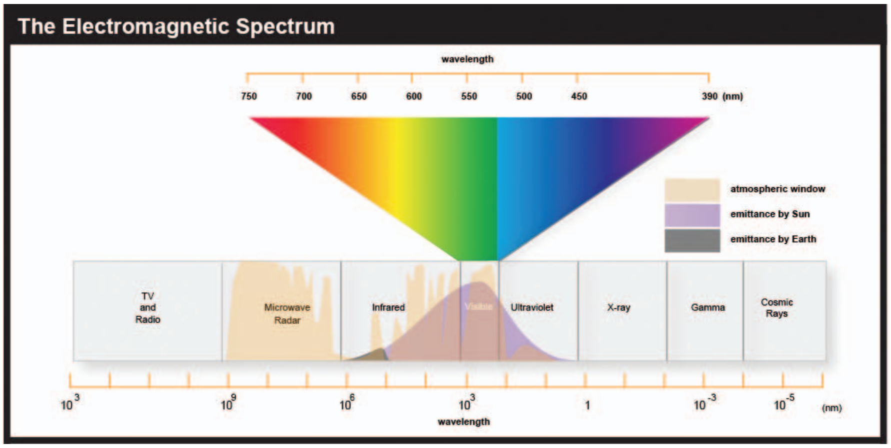

Figure 1 – Electromagnetic spectrum
Copyright Penn State's College of Earth and Mineral Sciences

What about the other way around? If our hearing range were restricted and only extended as far as our vision then we might be limited to 390 Hz – 700 Hz. In musical terms that's about G4 – F5, not quite one octave. What would that mean? There's the obvious that we wouldn't be able to hear bird song at one end or thunder at the other. But perhaps more interestingly, because that range is less than a single octave, we likely wouldn't have much concept of harmony or music. No idea that some sounds seem to belong together, while others grate. Within our limited range, we would likely be much more discriminatory—perhaps giving different wavelengths specific names, just as we do with colors in vision. We might then associate different colors/wavelengths of sound with specific moods and emotions, again just as we do with vision. Note that I'm talking about a specific single tone; you might argue that we hear emotion in sound now, but I might argue that's through harmony and combinations of sounds, not single notes.

Okay Mike, you are now saying, all this is perhaps mildly interesting, but not that exciting. I agree, however, stick with me please, I promise it gets better.

I'd argue that the real difference between hearing and vision is discrimination. Our vision system is rudimentary in terms of frequency discrimination. We have just three different receptors, each responsive to a band of wavelengths that we call colors. Worse, each of those three receptors is broad and they overlap. On the face of it, you couldn't design a poorer system for discriminating wavelengths!

Figure 2 shows the responses of the three human eye color receptors. You may have seen these described as long, medium, and short, or even red, green, and blue, but both sets of names over-simplify the situation. In reality, the response curves do not correspond to individual colors and instead overlap significantly. In particular, the long and medium cone responses overlap almost completely and the peak sensitivity of what we might think of as the red or long receptor falls in the area we call yellow.

There is a key concept embedded in the spectral response diagram that is critical to an understanding of color vision. Each cone type only has the ability to indicate that it has received a photon of light with a wavelength that falls somewhere within its acceptance band and the intensity of that input. However, it cannot tell us precisely where in that band the light falls. For example, the medium (green) receptor will fire for photons that have wavelengths anywhere between 400 nm and 700 nm. The output signal from that receptor to the brain will be identical for light anywhere in that range, and contains no information at all about the specific wavelength of that light. As illustrated in **Figure 3**, the photoreceptor signal when triggered by a light source at wavelength A would be identical to that for one of the same intensity at wavelength B.

Similarly, the same increase in output of a receptor could be triggered by either a higher intensity of the same wavelength, or by a different wavelength which happens to be closer to the peak sensitivity of that particular receptor. **Figure 4** shows that an increase in output of the same photoreceptor shown in **Figure 3** from light

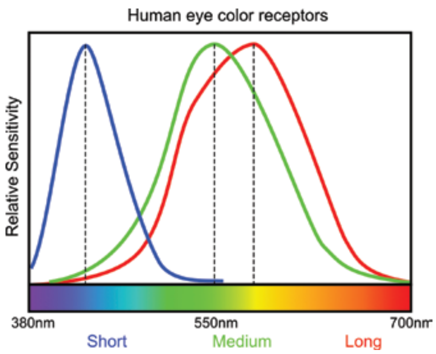

Figure 2 – Human eye color receptors

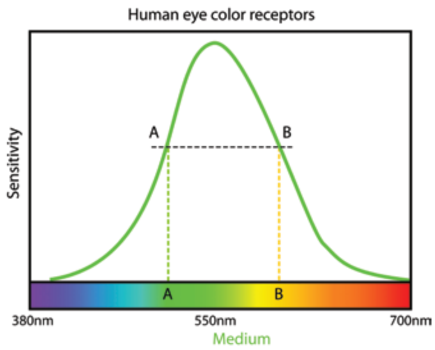

Figure 3 – Univariance – Wavelength change

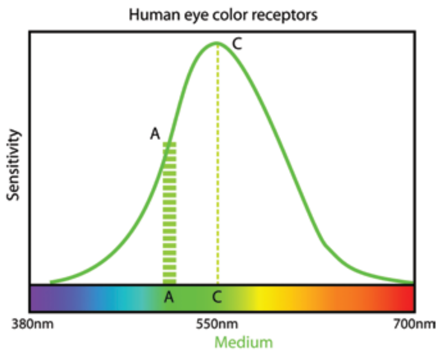

Figure 4 – Univariance – Intensity change

at wavelength A could either be triggered by an increase in the intensity of that light at A, or by a shift in wavelength to wavelength C where the photoreceptor is more sensitive.

There is no way to distinguish from the output signal from a single photoreceptor between a change in wavelength and a change in intensity. This principle, known as univariance, where a sensor has a single output sensitive to more than one parameter, is critical in understanding why the human brain is so important in interpreting the data. It is only by looking at the outputs of all photoreceptors in combination that determining color is possible.

Our hearing system, on the other hand, has no such restrictions. Instead the cochlea (shown unrolled in **Figure 5**) has something like 15,000 hair cells connected to over 30,000 nerve fibers and is capable of discriminating thousands of different frequencies individually. It appears that, for moderate loudness levels, humans can detect a frequency change of about 1 to 3 Hz for frequencies up to about 1,000 Hz. The fundamental principle of aural perception, Ohm's Law, (yes, it is the same Georg Ohm who is better known for his electrical law) is that the fundamental sound sensation corresponds to a simple harmonic vibration. **Figure 5** shows the cochlea unrolled from its usual snail-like spiral with the central basilar membrane and rows of hair cells. Low frequencies are detected at the larger end, higher frequencies at the narrow end. The blue line shows where maximal triggering might occur for a single tone range.

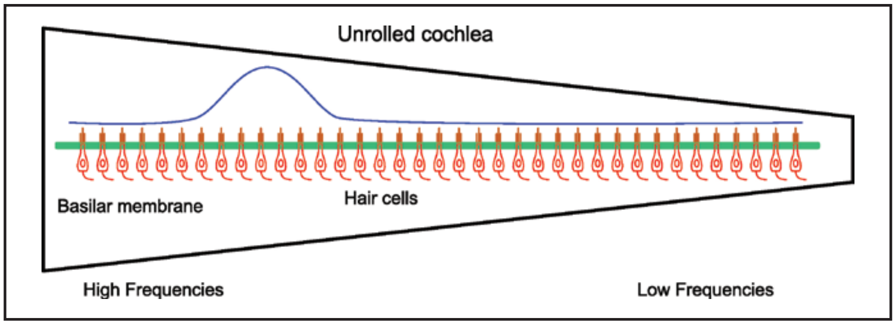

Figure 5 – Cochlea structure (unrolled)

Ohm's Law means that the separate simple frequency components can be recognized in a complex note. Later experiments have confirmed Ohm's Law, but without careful training by the listener it may not always be recognized.

I'm simplifying this considerably for the sake of clarity and to make my point. To paraphrase from an article by Jeffrey Hass at the Center for Electronic and Computer Music, School of Music Indiana University, Bloomington, IN. "When we listen to an orchestra, we might hear numerous instruments playing approximately the same pitch, with some very slight differences in frequency (called chorusing) that give the sound a depth and richness beyond what a single instrument would produce. A slight difference in frequency will lead to the phenomenon of beating. We perceive the slightly mistuned notes as a single chorused pitch up to the limit of discrimination, a difference of approximately 10 to 15 Hz, beyond which we hear two separate tones. At the very point of such a perceptual separation lies an area of tonal roughness. While a single pitch may maximally stimulate a specific spot on the basilar membrane in the cochlea, it also stimulates some adjacent hair cells as well. These lie within what is called the critical band. Other pitches which are close in frequency may also share some hair cells in common, which is theorized to cause intervalically close tones to sound more complex that more widely separated tones. The intervallic width of the critical band varies with register, being a large percentage of the frequencies of two low tones, and a smaller percentage of the frequencies at a higher register (it is about a minor third above A440). This may account for our orchestrational penchant for using wider intervals in lower registers. Another applicable phenomenon is that when two sounds of equal loudness are close in pitch, thereby in the critical band, their combined loudness will be only slightly greater than one of them alone."

The systems are thus fundamentally different. Vision has just three broadband frequency sensors and we use the combined output of those to perceive wavelengths that we call color. The key point here is that, whichever sensors of the three are triggered by whatever combination of wavelengths, we see it as a single result, a single color. Hearing in contrast has thousands of narrow band sensors and they are independent all the way to the brain. No combination of those sensors produces a single result. Instead we can recognize that multiple frequencies are present at the same time, and distinguish between them.

If vision were like hearing, then our color system would be entirely different. As it is now, if an object emits light in both the red and green wavelengths we see it as a single combined color, yellow. However, if our vision behaved like hearing then that wouldn't be the case. Instead we would see both the red and green colors independently and separately both at the same location. It's very hard for us to imagine what that might look like. There would be no such thing as metamers and no color mixing as we know it. If that sounds limiting, it isn't, it's just different. We would still have the ability to accurately distinguish between thousands of colors, as we do now, but each of those colors would consist of a single wavelength. They would be pure colors, not the mixed variety we are familiar with.

Some colors wouldn't exist any longer. In particular, the colors along the bottom, imaginary, edge of the normal color space diagram where blue loops round to red giving us a range of magentas. Magenta isn't a real color, it only exists as a mixture of red and blue, and our new hearing-style vision

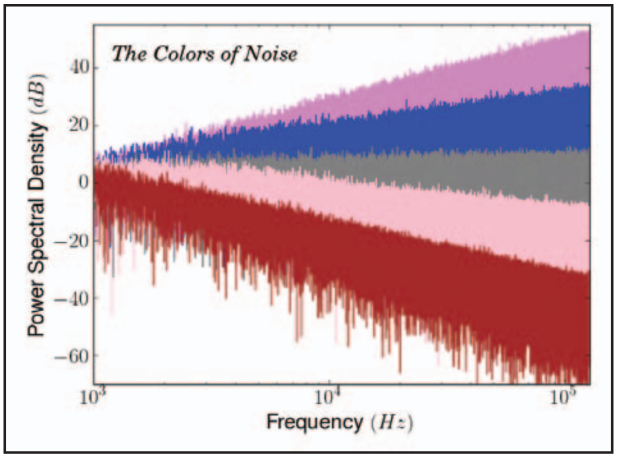

Figure 6 – The Colors of noise

system would see it as just that, red and blue together. Changing the relative intensities of the red and blue wouldn't change the hue, as it does now where we perceive a change from a blueish magenta to a reddish one, instead we would see the red and blue components increasing and decreasing independently. Similarly a pale pink might be seen as white, perhaps a broad band background signal like white noise, with a red on top of it. I can't show you a picture of all this, you have to treat it as a thought experiment and imagine it.

We might get a completely new effect with vision, that of true color harmony. Perhaps colors with related frequencies would blend together. In hearing, when two notes are separated by a musical third the result is perceived as a warm, close connection between the two notes. This musical interval is a mainstay of Western Music. An interval of a third is a specific ratio between the two notes, that of 5:4. The vision equivalent might be between, say, a red at 650 nm and a green at 520 nm, which have the same 5:4 ratio. Those two colors shown together might elicit a similar warm emotional response as the musical equivalent.

In both cases, we would still be able to distinguish between the two notes playing even though we are simultaneously conscious of the combined effect. When you play two notes on a piano, you hear the two notes separately as well as perceiving the effect of the combination. They don't merge into a new single note. However, if our hearing were like our vision, they would. This is very weird to imagine. Playing two notes to a vision-like hearing system would result in a single new note somewhere in between them. A simultaneous high pitched sound and low pitched sound would result in us hearing a single middle pitched sound. The frequency of that sound would vary depending on the relative loudness of the two original sounds. Playing a middle C and the C an octave above would result in the perception of an imaginary single note in the octave between them. Then, if we altered the relative intensities of those two notes you would be able to simulate any note in that range. Effectively sliding up and down the scale as the intensities changed. At no time would you be able to tell that there were two notes playing, it would always sound like a single note.

Sound additive mixing and sound metamers would exist. You could make a middle C by playing a combination of other notes with intensities adjusted to produce the perception of middle C. Just like an RGB color mixing luminaire, you could have an HML (High-Middle-Low) loudspeaker capable, as the manufacturer's advertising would say, of producing 16 million different notes! But, as with color in lights, only one at a time.

I've just scratched the surface of this analogy, and it is, as I unashamedly admit, a completely useless exercise. Except, perhaps, that it points up the special qualities of our vision and hearing systems that help us perceive the world in the way we do and makes us think about the why and the wherefore.

My musings took me further, into considering what were the vision analogies of white noise, pink noise, and all the other flavors of noise? What are the sight analogies of melody, or rhythm? What does subtractive color mixing sound like? Is it just filtering? What does timbre mean to a vision system? How about the perceptual differences between a square wave, a triangular wave, and a sine wave, very different aurally, but how about visually? Lots of things to waste time over. I'd love to hear any thoughts you have on other cross overs between sound and vision.

One final thing. In checking the facts on this article I found data that suggests that the smallest movement of the ear drum (tympanic membrane) that our hearing system is capable of detecting is 0.5 nm. Yes, I really mean nanometers, not millimeters. To put that in perspective, the mechanical system in our ear can detect motions that are 1,000 times smaller than the wavelength of the light our eyes see. That's just astounding. ■

Mike Wood runs Mike Wood Consulting LLC, which provides consulting support to companies within the entertainment industry on product design, technology strategy, R&D, standards, and Intellectual Property. A 40-year veteran of the entertainment technology industry, Mike is a past President of ESTA and Co-Chair of the Technical Standards Council. Mike can be reached at mike@mikewoodconsulting.com.

Note from the Editor:

It is difficult to comprehend, but this is Mike's 50th "Out of the Wood" column for our journal, *Protocol*. Thank you, Mike, for sharing the latest cutting-edge technologies and explaining the details in descriptive language readers can grasp and apply. Your *Protocol* column is a gift to our ESTA members and readers. We value your contributions and support—here's to 50 more! ~ The Editor