



Tasting music like wine: Sensory evaluation of concert halls

Tapio Lokki

JUKKA PÄTÄNEN

How do acoustics affect a concertgoer's experience? With the right tools, we can learn a lot by asking listeners to tell us in their own words.

When the intermission started, my wife was so excited she could hardly wait to get to the foyer. Over a glass of wine, she couldn't stop telling me how great the music we just heard was and how incredibly well she heard the pianissimos of the violin soloist. I was astonished: The performance had not touched me at all, even though the music was the violin concerto of Jean Sibelius, my favorite. The soloist had played well, but for some reason I did not feel the music and had not managed to enjoy myself. Had the acoustics of the hall ruined the otherwise good performance for me?

We got our wine, and it saved my evening. The first draft was rich and full bodied with a long-lasting aftertaste. Finally, I could concentrate on the good wine and forget the disappointing musical experience. My wife kept praising the performance and the acoustics of the hall—even declaring that the pale character of the wine could not take away her feeling of joy. Clearly, we perceived both the wine and the music differently.

At that moment, about six years ago, I realized that wine and concert-hall acoustics have a lot in common. They're each characterized by a multi-dimensional array of perceptual attributes. Evaluat-

ing them involves matters of personal taste. When comparing several acoustics or wines, different people concentrate on different aspects of the sound or taste, and they verbalize their perceptions differently. And yet winemakers have developed methods to determine what makes one wine better, worse, or different than another. The aroma wheel, for example, is a detailed characterization of the many flavors and fragrances found in wines. How has the wine industry been able to see past the large perceptual differences between individual tasters to understand the underlying characteristics that contribute to the overall quality of wine? And could their methods be tailored to the perceptual evaluation of concert-hall acoustics to better understand the multifaceted experience of a concert audience?

Concert-hall acoustics have been investigated since the pioneering work of Wallace Sabine more than a century ago,¹ and scientists have tried to understand which perceptual attributes contribute to the general opinion of extraordinary acoustics. To

Tapio Lokki (tapio.lokki@aalto.fi) is an associate professor in the department of media technology at the Aalto University School of Science in Espoo, Finland.

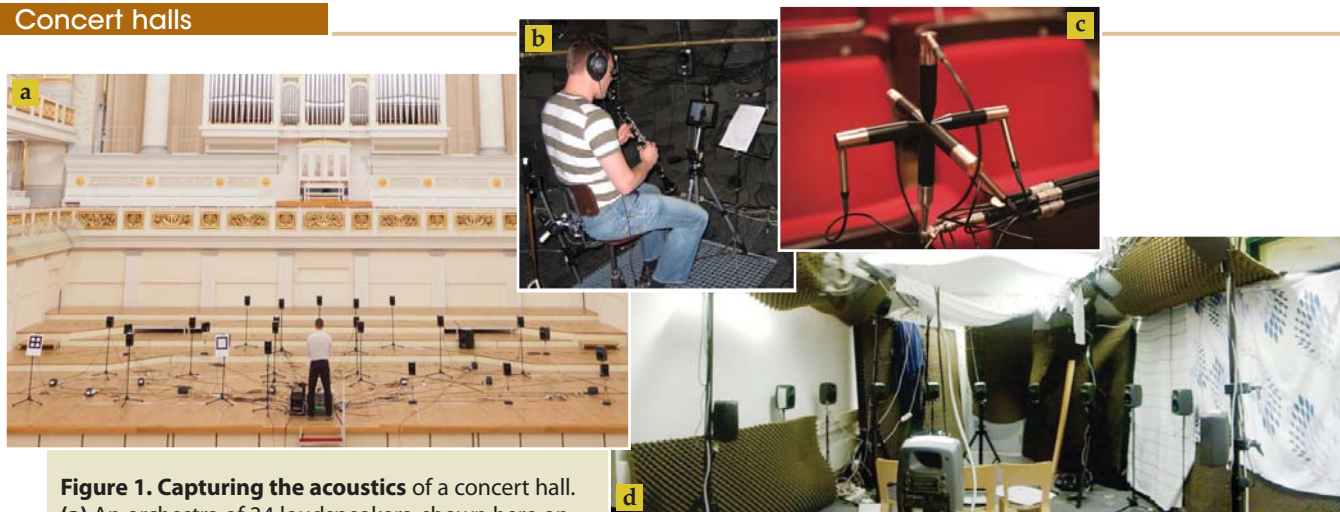


Figure 1. Capturing the acoustics of a concert hall. **(a)** An orchestra of 34 loudspeakers, shown here on the stage of the Konzerthaus Berlin, can reproduce the spatial sound output of a real symphony orchestra. **(b)** The instruments are recorded one at a time in an anechoic room. **(c)** An array of six microphones is placed in a seat in the audience. Because the complex spatial sound field produced by the full loudspeaker orchestra is too difficult to deconstruct, we record spatial impulse responses one at a time and convolve them with the instrumental recordings in the laboratory. **(d)** Twenty-four loudspeakers surround the listener in our listening room.

understand human response to the complex sound field in an enclosed space, research on room acoustics has applied both objective and subjective methods, similar to the wine industry.

Objectively, concert halls are studied by analyzing energy decays of sound. In particular, researchers measure impulse responses—that is, how a brief, simple acoustic signal broadcast from the stage is received as a function of time at some point in the audience. From those measurements, one can derive a number of acoustic parameters, such as reverberation time, sound strength, balance between early and late arriving sound energy, and measures of the spatial properties of the sound, as specified by the International Organization for Standardization.² The ISO standard proposes that the acoustics of a hall can be described with just a few numbers obtained by spatially averaging over several measured seats. That standard has recently been criticized on many grounds: The algorithms to compute the parameters are imprecise, the applied frequency range is too narrow, and a single omnidirectional source is a poor representation of the dozens of sound sources present in a real orchestra.³ Moreover, the objectively measured parameters fail to describe the details of perceived acoustics.

Subjective comparison of concert halls is not an easy task either. The music, the conductor, and the performance of the orchestra all affect the listening experience, and the contribution of the auditorium itself is hard to isolate with subjective surveys. Traditionally, those surveys have been conducted by distributing questionnaires to listeners at live concerts. Other research has involved interviews with conductors, musicians, and audience members. After 50 years of subjective research experience, Leo Beranek has developed a ranking of the best concert

halls in the world.⁴ Other significant studies have been done by research teams in Göttingen and Berlin⁵ and at the University of Bath.⁶ Subjective evaluations are also made under laboratory conditions with virtual acoustics techniques, which are based on convolving music signals with impulse responses, either captured from real halls⁷ or simulated via room acoustics modeling.⁸

Comparing samples

The fundamental requirement of the sensory evaluation techniques used in wine tasting is that the assessors are able to compare samples.⁹ Tasters are presented with a line of glasses filled with different wines, so they can taste wines in any order as many times as they want. After each glass they verbalize what they taste and create a vocabulary to describe the perceptual differences between the wines. Sensory evaluation can be performed using consensus vocabulary profiling, in which tasters first elicit adjectives to describe the wines and then discuss their assessments with other tasters in the group to develop a common set of consensus attributes. In individual vocabulary profiling, in contrast, tasters work independently. Each identifies several attributes that distinguish the wines—sweetness, for example, or muskiness—and ranks the wines according to each of those attributes. Statistical analysis of all the tasters' rankings reveals the most important differences among the wines, as perceived by the group.

Consensus vocabulary profiling requires assessors to be at least somewhat experienced so that they have a common understanding of the complex meanings of the words they use. Individual vocabulary profiling overcomes that need: It doesn't matter if one taster describes a wine as "fruity" and another describes it as "berrylike." If they rank the wines, from most to least fruity or berrylike, in nearly the same way, it's likely that they've identified the same attribute. Sometimes the individual perceptions can be quite different, but with a reasonable number of assessors—15 or more, in practice—common salient characteristics can be found.

The obvious challenge in applying sensory evaluation to concert-hall acoustics is the requirement of simultaneous comparison of halls. The human auditory memory is too short for listeners to reliably compare concert halls by listening to music

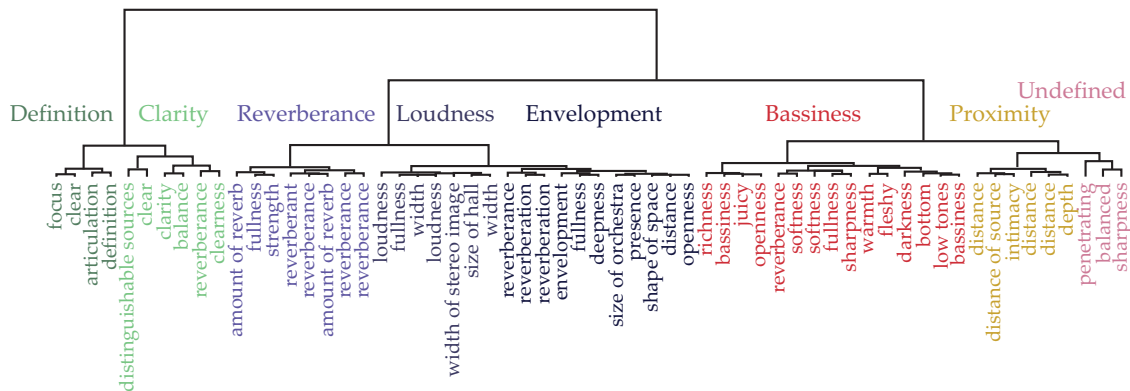


Figure 2. The hierarchical clustering of 60 attributes elicited and rated by 17 assessors in a study of nine concert halls. The study was conducted in Finnish with Finnish listeners; the words and phrases shown here are translations. Each word or phrase represents a single attribute named by a single listener—so, for example, the fact that “sharpness” appears twice means that 2 of the 17 assessors used that word to identify one of their attributes. The vertical dimension represents a measure of the difference between attributes or clusters of attributes. (Adapted from ref. 16.)

in situ in different halls. The acoustics of different concert halls—and of different seats in those halls—must somehow be recorded and reproduced in the laboratory so that listeners can switch between different acoustics in the blink of an eye. Moreover, the recorded music must be played by the same orchestra at exactly the same level and tempo in each hall. Even a professional orchestra cannot do that, because musicians intuitively adjust their playing style according to the acoustics. To isolate the acoustics as the only variable in the recordings, my research group’s solution was to build an orchestra out of loudspeakers, as shown in figure 1a.

The loudspeaker orchestra consists of 34 calibrated loudspeakers arranged on the stage in the form of a real orchestra. The loudspeakers’ placement and direction are designed to match those of real musical instruments as well as possible. For example, each violin channel consists of two loudspeakers, one pointing toward the audience and another lying on stage pointing up because violins emit most of their sound into the upper hemisphere, particularly at high frequencies.¹⁰

Each loudspeaker on stage has to reproduce the sounds of its respective instrument as cleanly as possible, uncontaminated by any other instrument or by the acoustic response of the room in which the instrument was recorded. Our solution was to record musicians one at a time in a small anechoic room.¹¹ While playing their own part, musicians listened to a piano track of the music and watched a video of the conductor on a small screen, as shown in figure 1b. Surprisingly, the orchestral musicians were able to play in tune and in tempo, even without any visual or aural cues from surrounding musicians. We recorded four excerpts of symphony music, 2–4 minutes each, from different periods and with different-sized orchestras.

When we played the recorded music through the loudspeaker orchestra in a concert hall, it sounded very realistic. But to record the spatial sound experienced by the audience in the hall and reproduce it in the laboratory, we would have to record the direction of incidence of different parts of

the complex sound field. That proved to be an insurmountable challenge. Plenty of techniques have been developed for spatial sound recording and reproduction,¹² but their goal is to make artistically good recordings, not authentic reproductions of sound.

Instead, we performed spatial sound recordings via impulse responses. From each of the orchestra’s 34 loudspeakers, one at a time, we played a logarithmic frequency sweep, and we measured the response with an array of six omnidirectional microphones arranged in a single seat in the audience, as shown in figure 1c. Those impulse responses can be analyzed to estimate the direction of incidence of the sound energy as a function of time and frequency.¹³ By convolving the anechoic instrument recordings with the spatial impulse response data, we calculated the signals we needed to distribute through the 24 loudspeakers in our listening room (shown in figure 1d) in order to authentically reproduce the

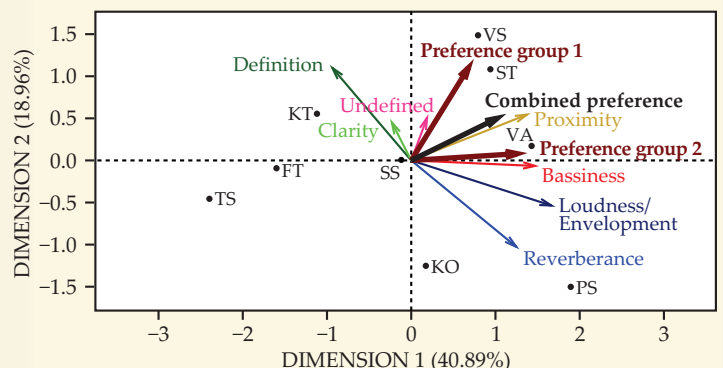


Figure 3. A hierarchical multiple-factor analysis of the attributes shown in figure 2 and the listeners’ preference ratings of the same halls showed that nearly 60% of the variance in the data was captured by the first two principal components, plotted here. The black dots and two-letter codes represent the nine halls themselves; the colored arrows represent clusters of attributes or preferences. The listeners’ preferences fell into two distinct clusters, shown in brown. Curiously, the average preference of both groups, shown by the black arrow, is well correlated with the perceived proximity of the orchestra. (Adapted from ref. 16.)

spatial sound of the loudspeaker orchestra playing in the measured concert hall.¹⁴

Attributes and preferences

So far, we have recorded 20 European concert halls for sensory evaluation. Because the orchestra and recording system are calibrated, the only variable in the samples is the acoustics, determined by the architecture of the concert hall. Thus the requirement of immediate comparison, similar to having a line of wine glasses, is fulfilled, and we are able to study concert-hall acoustics in great detail.

Sensory evaluation works remarkably well for assessing the acoustic differences between concert halls and between seats in one hall. In particular, we've found individual vocabulary profiling to be a reliable method for extracting perceptual differences between concert halls. First, the words used by the listeners give us a rich vocabulary to understand the salient differences between halls. Second, the clustering of those attributes in multidimensional space reveals consensus attribute groups and sensory profiles of studied concert halls. Third, the studied concert halls can be ranked according to each of those attribute clusters. If we ask the listeners to also rank the halls in order of preference, their collective preferences can be related to the sensory profiles of the halls.

In our first major study, we asked 20 listeners to evaluate three recording positions in each of three Finnish concert halls.¹⁵ Some recorded seats were in the balcony, and others were close to the orchestra. The assessors elicited and identified a total of 102 attributes. But our analysis revealed that just one cluster of attributes, related to overall volume and perceived distance, explained more than 50% of the variance in the collected data. The result is unsurprising because the physical distance of recording positions varied a lot. Less obvious was our finding that different listeners using the same word sometimes mean entirely different things. For example, attributes described by different listeners as “reverberance” fell into two distinct groups. Some were clustered with attributes described by other listeners as having to do with the perceived size of the space. Others were clustered with attributes related

to envelopment—sound arriving from all directions, not just from the front. Such a finding would not have been possible in a listening test with the attributes defined by the researchers. If we'd asked all the listeners to evaluate the halls according to their reverberance, we would not have had the tools to find out what they thought the word meant. The sensory evaluation and individual vocabularies showed their power in providing rich perceptual data.

For the next study, the physical recording distance was fixed to 12 m in each concert hall and only one seat from each of nine halls formed the stimulus set.¹⁶ Seventeen assessors offered a total of 60 discriminative attributes, all but 3 of which could be grouped into seven main clusters, as shown in figure 2. The largest group contains clusters of attributes related to loudness, envelopment, and reverberance. (In this case, “reverberance” was not clearly associated with either of the two meanings identified in the previous study.) The second large group comprises bassiness and proximity attributes. The third contains definition and clarity attributes. Similar discriminating factors have also been found by Beranek.⁴

We also had the listeners rank the nine halls in order of preference. Each ranking—by preference or by some discriminative attribute—can be represented by a point in a multidimensional space. By performing a hierarchical multiple-factor analysis on the data in that space, we found that the first two principal components explain almost 60% of the variance in the data. By projecting the space onto those two dimensions, as shown in figure 3, we can visualize most of the differences between attributes, preferences, and halls. Assessors could be divided into two groups based on their preferences. The first group preferred concert halls that render proximate sound with high definition and clarity. In other words, they liked relatively intimate sound in which they could easily distinguish individual instruments and melody lines. The second group preferred a louder and more reverberant sound with good envelopment and strong bass. All assessors disliked the concert halls with weak and distant sound. Surprisingly, the best correlation with average preference ratings of both was perceived proximity.

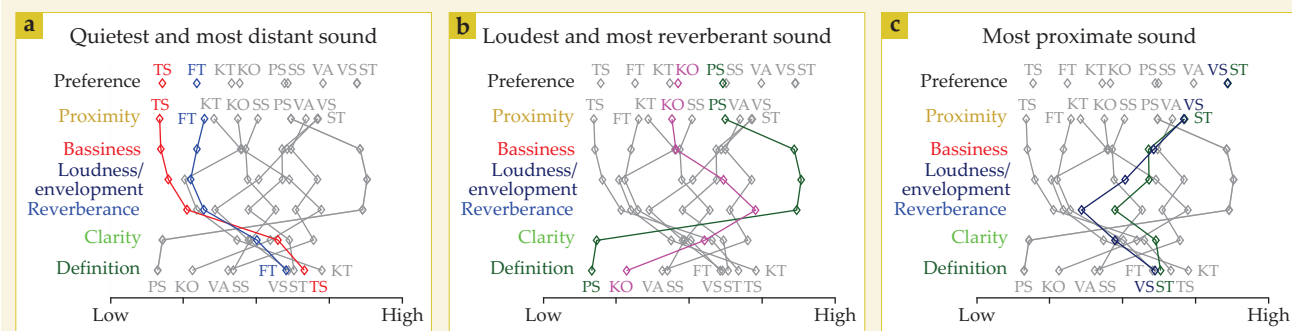


Figure 4. Sensory profiles, based on the data shown in figures 2 and 3, of differently perceived concert halls. **(a)** Halls TS and FT produced quiet, distant sound and were rated poorly by the listeners. **(b)** Halls PS and KO, with their loud, reverberant sound, were preferred by some listeners but not all. **(c)** The favorite halls, VS and ST, were the ones that offered the greatest subjective proximity. Sound samples for these six halls are available with the online version of this article.

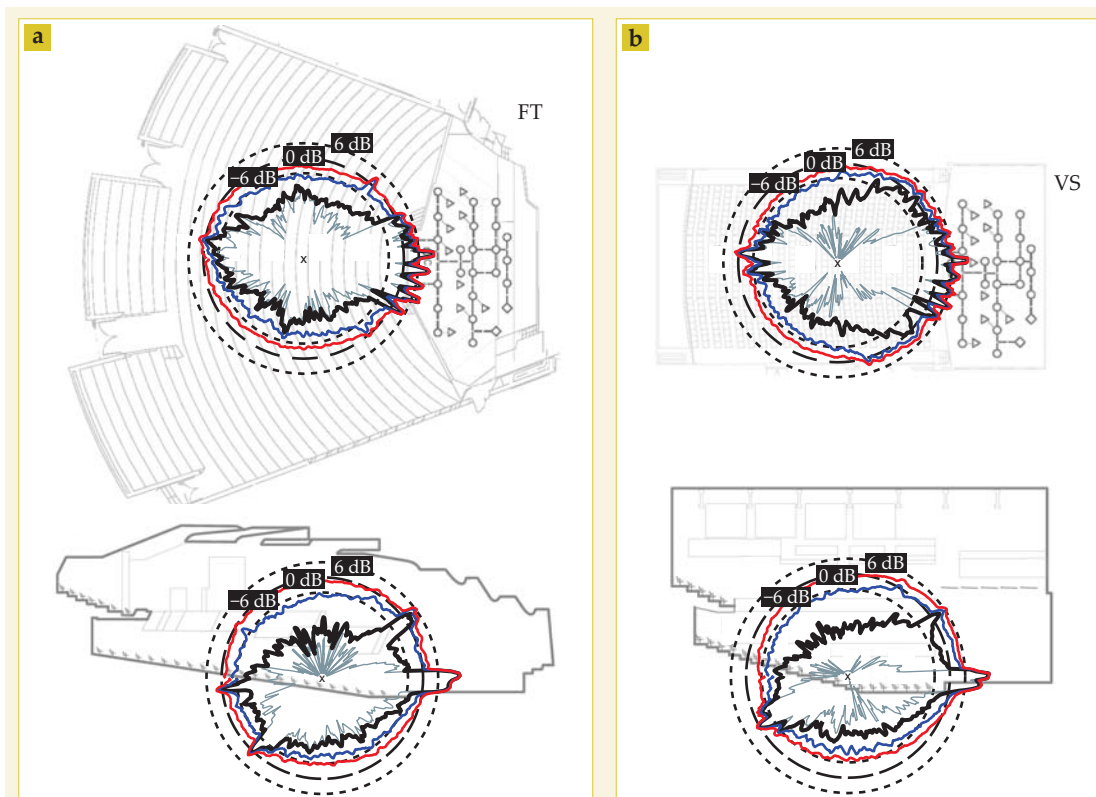


Figure 5. Visualizations of sound energy received by a listener in a concert hall as a function of direction and time. The jagged curves represent the cumulative sound energy received within 5 ms (light gray), 30 ms (black), 100 ms (blue), and 2000 ms (red) after the arrival of the direct sound. The plots are superimposed on top and side views of the hall architecture to show the relationship between hall design and acoustics. **(a)** In hall FT, the fan-shaped design and low ceiling yield a strong early reflection from above but little sound reflected directly from the side walls. Both of those characteristics contribute to the perception of a distant sound, disliked by the listeners. **(b)** In hall VS, one of the most preferred halls, the situation was just the opposite.

The same data can be represented with sensory profiles, as plotted in figure 4. The most disliked halls, highlighted in figure 4a, have similar profiles with quiet, distant sound. Interestingly, the loudest and most reverberant halls, as shown in figure 4b, offered poor clarity and definition, so they did not render the most intimate sound. They lie in the middle of the average preference ranking. The most preferred halls, highlighted in figure 4c, have enough loudness and envelopment, good definition, and bass that contribute to intimate sound.

The loudspeaker orchestra was recorded at a distance of exactly 12 m in each hall. Why, then, does the perceived distance vary so much from hall to hall? We recently developed a technique for spatiotemporal visualization of cumulative sound energy that might help to explain the difference.¹⁴ Using the same spatial impulse response data that we measured with the loudspeaker orchestra in each hall, we plot the arrival of sound energy as a function of both direction and time. Overlaying the plots on architectural drawings of the hall helps us to identify reflecting surfaces that cause sound to arrive from directions other than from the front of the hall.

Figure 5 shows spatiotemporal plots for two of the halls we studied. Concert hall FT, with its fan-shaped architecture, rendered a perceptually distant sound and was one of the least preferred halls.

Concert hall VS, on the other hand, rendered a sound that was perceived as most proximate and was among the listeners' favorites. The thick black lines in figure 5 represent the cumulative sound energy arriving within 30 ms after the direct sound. They show that in hall FT, most of the early sound arrives from the forward direction, whereas hall VS features strong early reflections from the sides. Moreover, hall FT has a prominent early reflection from its low ceiling, and hall VS has only a modest early vertical reflection (which actually arises from reflectors above the stage, not from the ceiling). We find that those two halls are representative of a larger pattern: Early lateral reflections contribute to a more intimate sound, whereas a strong early ceiling reflection causes the sound source to be perceived as more distant.¹⁷ The reason is that the reflections from the side are amplified more than those from above, in particular at high frequencies, due to the shape of the human head.¹⁸ (See the article by Bill Hartmann in *PHYSICS TODAY*, November 1999, page 24.)

Conclusions

Sensory evaluation methods, borrowed from the food and wine industry, are useful for studying concert-hall acoustics because they can extract information often hidden behind preference judgments.

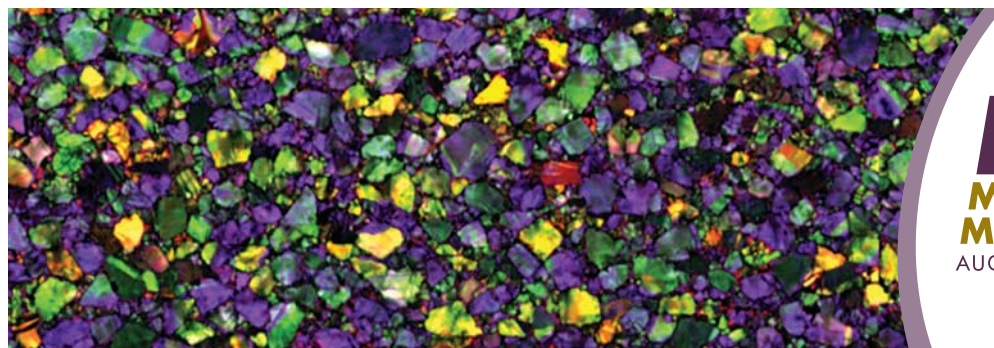
With such methods, in particular those based on individually elicited attributes, one can develop sensory profiles of concert halls or of seats inside one concert hall. Preference judgments might give an overall average picture, but the variance in the data is typically large due to the assessors' personal tastes and previous experiences. Sensory evaluation methods provide a link between those subjective preferences and perceptual characteristics. Our research is zeroing in on the main characteristics that form the basis of that multidimensional perceptual space.

Sensory evaluation requires immediate comparison of studied samples, whether they are wines or concert halls. That requirement has led us to develop a symphony orchestra simulator that has become a valuable research tool. We can now listen to an authentic reproduction of music as experienced in a concert hall, and we can analyze the spatial characteristics of the sound field in the measurement position. In the near future, those tools will pave the way for a comprehensive understanding of the links between architecture, music, acoustics, and human perception. My wife and I will soon understand why we perceive music in our local concert hall so differently and what features of the architecture influence our perceptions.

I thank all the researchers on my Virtual Acoustics team for their hard work, and the European Research Council (203636) and the Academy of Finland (257099) for financial support.

References

1. W. C. Sabine, *Collected Papers on Acoustics*, Harvard U. Press, Cambridge, MA (1922), p. 3.
2. International Organization for Standardization, ISO 3382-1: 2009, "Acoustics—measurement of room acoustic parameters—part 1: Performance spaces," ISO, Geneva (2009).
3. J. Bradley, *Appl. Acoust.* **72**, 713 (2011); L. Kirkegaard, T. Gulsrud, *Acoust. Today* **7**, 7 (2011).
4. L. Beranek, *Concert Halls and Opera Houses: Music, Acoustics, and Architecture*, 2nd ed., Springer, New York (2004).
5. L. Cremer, H. Müller, *Principles and Applications of Room Acoustics*, vol. 1, T. J. Schultz, trans., Applied Science, New York (1982), chap. 3.
6. M. Barron, *Acta Acust. United Acust.* **66**, 1 (1988).
7. G. Soulodre, J. Bradley, *J. Acoust. Soc. Am.* **98**, 294 (1995).
8. J. Bradley, G. Soulodre, *J. Acoust. Soc. Am.* **98**, 2590 (1995).
9. H. Lawless, H. Heymann, *Sensory Evaluation of Food: Principles and Practices*, Aspen, Gaithersburg, MD (1999).
10. J. Pätynen, T. Lokki, *Acta Acust. United Acust.* **96**, 138 (2010).
11. J. Pätynen, V. Pulkki, T. Lokki, *Acta Acust. United Acust.* **94**, 856 (2008).
12. F. Rumsey, *Spatial Audio*, Focal Press, Boston (2001).
13. S. Tervo et al., *J. Audio Eng. Soc.* **61**, 17 (2013).
14. J. Pätynen, S. Tervo, T. Lokki, *J. Acoust. Soc. Am.* **133**, 842 (2013).
15. T. Lokki et al., *J. Acoust. Soc. Am.* **130**, 835 (2011).
16. T. Lokki et al., *J. Acoust. Soc. Am.* **132**, 3148 (2012).
17. T. Lokki et al., *J. Acoust. Soc. Am.* **129**, EL223 (2011).
18. T. Lokki, J. Pätynen, *J. Acoust. Soc. Am.* **130**, EL345 (2011).



4 FULL DAYS Symposia, tutorials, exhibits, posters, short courses & workshops

Visit Often for Updates!

<http://microscopy.org/MandM/2014>

Call for Papers

Paper Submission Open: November 15, 2013

Paper Submission Portal

Paper Submission Deadline: February 8, 2014

Program Details

Awards Information

Hartford Area & Hotel Information

Hotel reservations available January 2014

Online Registration

Opens February 2014

Image at Top: Cathodoluminescence image of composite SiC-diamond cutting tool; Colin MacRae, CSIRO-Minerals, Clayton, Victoria, Australia

