Proceedings of the 1989 International Computer Music Conference, 268-272. San Francisco, CA: The Computer Music Association, 1989.

PERCEPTION OF CONCURRENT TIMBRES AND IMPLICATIONS FOR ORCHESTRATION

Gregory J. Sandell
Northwestern Computer Music
Northwestern University
Evanston, IL 60208
greg@music.nwu.edu

Candidates for perceptual correlates to timbre and orchestration are considered, and one of them (blend) is studied in an experiment using the Stanford collection of synthesized musical instruments. Results suggest that the blend judgment is inversely related to either the sum of the centroids of the two tones, or the sum of the attack times. The lower the overall centroid, or attack duration, the better the blend.

Most of the influential studies in timbre which have appeared over the last 15 years have cited Schoenberg's and Webern's revolutionary contribution to orchestration commonly known as *Klangfarbenmelodie*. The fact that composers have long been seeking to create timbral worlds in which discrete categories of timbre could be apprehended by a listener makes the historical comparison inevitable. Although this interest reflects the fact that orchestration is the musical context in which we most frequently encounter timbre, we may ask if *Klangfarbenmelodie* is the kind of orchestration in frequent practice. Indeed, with the multi-channel, multi-timbral resources available in today's MIDI workstations, many composers are "orchestrating" with fixed resources (presets) in the pre-Schoenbergian sense of the word. The question, however, is what sorts of studies in timbre perception would be informative to this type of orchestrational practice?

Orchestration as a synonym for instrumentation---effective ranges of instruments, the varying character of instruments' registers, the weaknesses or impossibilities of shakes, trills, double stops, and so on---is of little use. Another topic under the umbrella of orchestration is the building of textures, an area in which timbre is only one of many components, since texture is just as much a matter of rhythm and pitch. As usual, we will have to narrow a multidimensional topic somewhat in order to investigate it meaningfully. I submit that *concurrent timbre*----the effects produced by different simultaneous timbres---is a subtopic of orchestration which is interesting to musicians, and can be perceptually investigated. The process by which hybrid, synergistic timbres sum from two separate sources is largely a mystery; orchestration treatises tell us that X "blends well" with Y, but what we would really like is some acoustical/perceptual information that would lead to a systematic approach to orchestration.

Psychoacoustic Aspects of Orchestration

The objective of the current research, is to 1) find what kinds of evaluations listeners make, or what kinds of perceptual reactions they have, to concurrent timbre; 2) measure those sensations; 3) correlate them to physical aspects of the signal. Let us consider some of the key issues involved in this search.

Differential Salience of Instruments. In some cases the character of one instrument dominates over another although the tones of both can be equally heard; e.g., oboe-trumpet can sound more like two trumpets than oboe. Certain inequalities between features of each tone may lead to this condition: rate of change in temporal envelope, within-instrument inharmonicity, and pitch jitter are some of the possible acoustic inequalities. Cognitive reasons may include dominance due to the greater recognizability of one tone, or, since onsets are so critical for instrument recognition, informational masking caused by a noisier attack in one of the two.

Masking Among Harmonics. The hybrid timbre formed by two instruments which are unlike the solo properties of either may be due to masking patterns between the spectra of tones. A study by Pepinsky (1941), for example, investigated a brass quintet playing a widely-spaced Bb major chord, employing various combinations of dynamic markings among individual players; based on predictions from masking data and the instruments' spectral descriptions, he evaluated the (expected) resulting tone quality. Mostly we expect to see masking of upper harmonics from upward spread of lower, louder harmonics; additionally, there may be some masking between upper harmonics which fall in one critical band. However, most currently available data on masking concerns sine tones of fixed frequency and amplitude, while little is known about the masking patterns between temporally varying components. The synergistic effects frequently referred to in orchestration treatises (where one tone "absorbs," "modifies," or "softens" another) may be the outcome of more complex kinds of masking, such as suppression (Shannon, 1976) and comodulated masking release (Hall, 1987).

Blend. Where a combination of timbres is intended to transmit a single musical stream, it may be desirable to experience them as a single fused image rather than as two kinds of sounds. This is a recurring concern of orchestration treatises; for example, Rimsky-Korsakov (1911) devotes a chapter to recommended pairings of instruments in unisons, thirds, sixths and octaves, emphasizing the avoidance of combinations where "each can be heard separately" (p. 61).

Investigating Blend

The attribute of Blend suggests a psychological relation between two tones, which, like similarity, can be rated on a magnitude scale. This puts us in a position to correlate sensations to physical aspects of timbres. As a first step in investigating orchestration, a study of blend was conducted, and is described here.

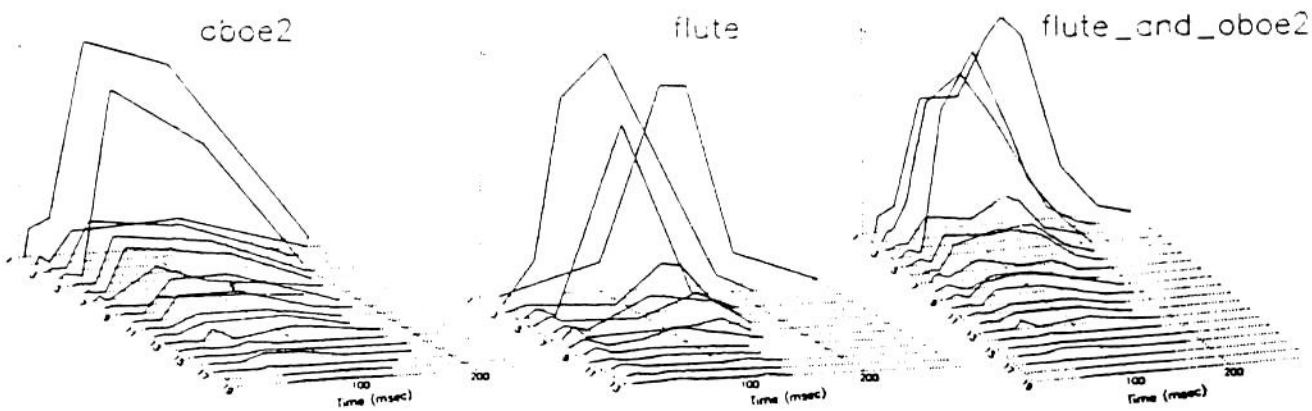

Figure 1. Time-amplitude-harmonic plots indicating the outcome of combined timbres.

Fifteen of the instruments used in Grey(1975), in the *line-segment approximations* form, were played in simultaneous presentation on a unison Eb4 to a group of eight musicians who rated their blend.[1] The original durations (ca. 330 msec) and relative intensities of the tones (which had been adjusted for equal loudness by Grey's listeners) were retained, while the pitch of the tones and the relative start times between pairs were adjusted to avoid distracting cues of pitch-beating and splattered attacks. Gordon's PAT (perceptual attack time) was used for the start-time adjustments (Gordon, 1984). A composite of one of the pairs is shown in figure 1. Subjects used a ten-point scale to rate blend, defined as a continuum between "fused" (10) and "separated" (1). Listeners rated all possible pairs (120, when including same-instrument pairs) exactly four times. All listeners showed significant agreement within their own four blocks of data, and all but one listener showed agreement between all other subjects. The data of the seven agreeing subjects was averaged together, and is the data which is discussed below.

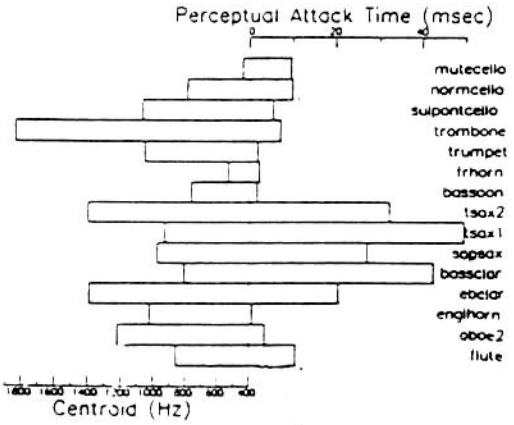

Figure 2.

Two aspects of the individual tones seem to determine the judgments. One is spectral centroid (Grey and Gordon, 1978), or the "bright-dark" dimension of the tones, and the other is the length of attack, here specified as the Perceptual Attack Time. First, a dark instrument shows higher (and more statistically reliable) blend judgments when paired with other dark instruments, and lower blend judgments as its pair increases with brightness. However, the presence of any dark instrument in a pair generally leads to a higher blend judgment than any pairing of two bright instruments. The effect may be observed in Figure 3, where a single point on this graph, say, the muted cello, represents the average rating for all pairs of instruments which contained the muted cello. Similarly, if one of the instruments in the pair has a quick attack, the pair will blend better than cases where both have a slow attack (fig. 4 shows the affect of attack characteristics in a way analogous to fig. 3). Here the correlation is not as strong but still significant. As a rule of thumb we may propose the following: the blend judgment is inversely related to either the sum of the centroids of the two tones, or the sum of the attack times. The lower the overall centroid, or attack duration, the better the blend. The trend is far from perfect, however, with several outliers which merit further investigation.

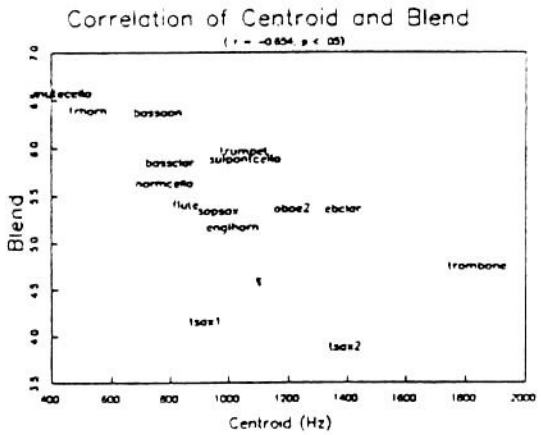

Figure 3.

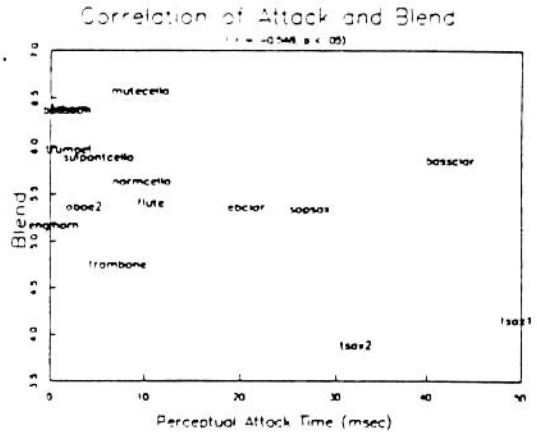

Figure 4.

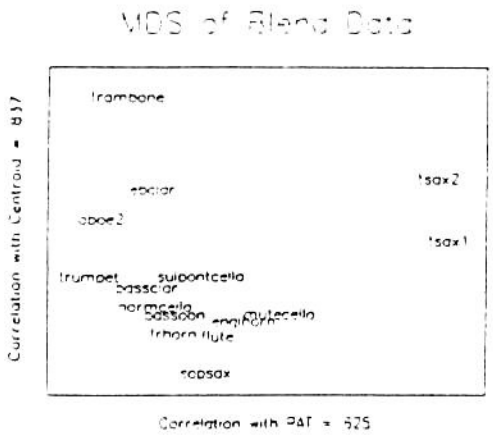

Figure 5.

A multi-dimensional scaling (MDS) analysis of the data, using blend ratings as distances, was conducted. The solution (figure 5) shows the blend between any pair of tones by their spatial proximity. Unfortunately, this analysis suffered from the fact that three instruments steadily got low blend ratings by the listeners: tsax1 and tsax2 (probably due to inharmonicity and pitch jitter) and the muted trombone (probably due to its extreme brightness). If the majority of Classic-Romantic orchestral instruments led subjects to make their judgments according to those cultural norms, it could explain the strong differences, at least, for the two saxophones. The result was a degenerate solution, where instruments were spaced excessively close to account for three anomalous instruments. However, correlations showed that two of the dimensions successfully recovered the centroid and attack dimensions.

Discussion

Other studies of timbre lend support to these findings in one way or another. Grey (1975) also showed centroid and attack to be primary determinants of similarity ratings for the same collection of tones. In fact, using Grey's MDS solution as a starting configuration for this study's MDS (such as it was) showed some match between the two solutions.[2] Harmonic synchrony, another dimension found by Grey, showed no apparent relationship to the current data; both synchrony within and across a pair of instruments was investigated. This is a curious finding as the temporal discontinuities introduced in the combination of instruments were considerable in some cases (see figure 1). Possibly the range of harmonic asynchrony in the collection of tones was not great enough to emerge as a factor in the face of the more prominent cues of spectrum and onset. Rasch(1978), investigating what leads to the perception of one vs. two notes when two are presented, showed that similarity of onset was the primary cue, but in cases of identical onset listeners used differences in spectral envelope as cues. A recent article by Belkin (1988) also advocates the view that spectral and temporal similarity determine the perception of blend. While their findings bear a resemblance to those of the present study, Rasch and Belkin do not show the bias for lower centroids and shorter attacks.

Orchestration treatises are the primary repository for attitudes and knowledge of timbre throughout the history of music, although they have rarely been considered as a source for contemporary thought on timbre. Their recommendations (as well as practices observable in scores) are corroborated by the current findings. The french horn and the bassoon (both dark instruments with quick attacks) are frequently called upon to combine with many instruments, because they blend so well with instruments having widely varying degrees of brightness and methods of production. The french horn is seen equally often paired with strings, woodwinds, and brass, while the bassoon is paired with the woodwinds, low strings or low brass. Brighter instruments, such as the oboe and trumpet, tend to be seen more as solo instruments or the top voice of their own section, rather than as elements of mixture with contrasting timbres. The fact that the centroid/attack net caught some well-known fish (bassoon-french horn) but some strange creatures as well (alto saxophone-cello) suggests that the orchestrational values specific to one musical era have more widely applicable perceptual underpinnings.

References

Belkin, A. 1988. "Orchestration, perception, and musical time: A composer's view." *Computer Music Journal* **12(2)**: 47-53.

Gordon, J.W. (1984) *Perception of Attack Transients in Musical Tones.* Report No. STAN-M-17, Center for Computer Research in Music and Acoustics, Stanford University.

Grey, J.M. 1975. *An exploration of musical timbre.* Report No. STAN-M-2, 1975, Center for Computer Research in Music and Acoustics, Stanford University.

Grey, J.M. and Gordon, J.W. (1978) "Perceptual effects of spectral modifications on musical timbres." *Journal of the Acoustical Society of America* **63**, 1493-1500.

Hall, J.W. 1987. "Experiments on comodulation masking release." In Yost, W.A. and Watson, C.S. (Eds.), *Auditory Processing of Complex Sounds.*, 57-66. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Pepinsky, A. 1941. "Masking effects in practical instrumentation and orchestration." *Journal of the Acoustical Society of America* **12**, 405-408.

Rasch, R.A. 1978. "The perception of simultaneous notes such as in polyphonic music." *Acustica* **40**, 21-33.

Rimsky-Korsakov, N. 1911. *Principles of Orchestration.* Trans. by E. Agate. New York: Dover, 1964.

Shannon, R.V. 1976."Two-tone unmasking and suppression in a forward-masking situation." *Journal of the Acoustical Society of America* **59**, 1460-1470.

---

1 I would like to thank David Wessel for his assistance with the stimuli.

2 These data points were figured from Grey(1975) by eye and a scanning device. The actual values for the solution are not given in his dissertation.

1989 International Computer Music Conference

Proceedings

The Ohio State University
Columbus, Ohio