M.I.T. Media Laboratory Perceptual Computing Section Technical Report No. 475
Appears in Proceedings of the XII Colloquium on Musical Informatics, Gorizia, Italy, September 1998

Analysis of Affective Musical Expression With the *Conductor's Jacket*

Teresa Marrin and Rosalind Picard
MIT Media Lab
20 Ames Street, E15-491
Cambridge, MA 02139
marrin, picard@media.mit.edu

Abstract
The *Conductor's Jacket* is a wearable physiological monitoring system that has been built into the clothing of an orchestral conductor; it was designed to provide a testbed for the study of emotional expression as it relates to musical performance. We used the *Conductor's Jacket* to gather and analyze data from a professional conductor in Boston during rehearsals of Prokofiev's *Romeo and Juliet Suite No.2*. Our findings indicate that several forms of expressive communication can be measured and detected in physiological signals. These include the use of *handedness* to emphasize musical changes, the signaling of upcoming events with sudden changes in effort, the difference between information-bearing and non-information-bearing gestures, the indication of intensity and loudness with changes in muscular force, and the use of breathing to express phrasing in the music.

Introduction
Recent work in the domain of affect recognition and physiological monitoring has yielded important results on the nature and expression of human emotions[5]. For example, several early studies have pointed to the presence of a 'contagion effect' whereby emotions can be transmitted from one person to another[2]. The presence of this effect explains why stress can be communicated between people under various conditions; it has also been hypothesized that other states can be transmitted contagiously. The precise mechanism through which this transmission occurs remains unknown, although we suspect that gestures and body language play a big role.

One promising new direction for the study of contagious emotional expression is in the performing arts, particularly in music. Music has often been described as a direct conduit for the communication of emotion; it might be said to be an ideal carrier channel for the transmission of affective information. Correspondingly, musical performers might be said to modulate the structure of musical scores in order to convey affecting and dramatic performances. The act of performing for an audience often requires the performer to project amplified or enhanced emotional states, and to this end performers often train for years to be able to effectively
and intentionally express these states. Several early and influential studies on emotional expression discuss this phenomenon with respect to performed classical music[1].

We chose to look at a very specialized form of musical performance, which is optimized for the transmission of emotional and dramatic expression: the role of the orchestral conductor. Conductors use a unique gestural language that combines both technical and affective information about a piece of music in real-time in order to aid those who are performing it. We hypothesize that conductors form expressive intentions for certain pieces that they then convey by means of gestures, and that the affective information is essentially encoded in the carrier signal of the beat-pattern. We hypothesized that the affective content of these signals might be decoded (as by musicians in an orchestra) by noting the difference between the conducted signals and the minimum amount of information that would have been required to execute an *unexpressive* (or minimally expressive) version of the same piece.

In order to test our hypotheses, we designed and built a system to robustly and unobtrusively sense expressive information from conductors under professional rehearsal conditions. This system had to be noiseless, light, not distracting or uncomfortable to wear for long periods of time, and able to withstand punishing conditions of extensive muscular activity, heat, and sweat. The resulting system, called the *Conductor's Jacket*, is a wearable network of physiological sensors that has been custom designed and embedded in clothing that is fitted to the wearer[4]. The jacket contains sensors for heart rate, respiration, skin conductance, temperature, and muscle tension. For muscle tension, we used four electromyogram (EMG) sensors, one on each bicep and tricep. These measure the small voltage created when the muscle generates force; the voltage is proportional to the instantaneous force output of the muscle. All of the sensors are held in place by elastic bands that have been sewn into the cloth of the jacket.

The data we collected supports three major features in the standard conducting technique: the use of the left hand to add emphasis and expressive information, the turning of pages so as to not attract attention or convey musical information, and the use of force in performing a beat gesture to indicate the volume and articulation with which that beat should be played. We also found some surprising results, including several instances where the muscles went limp right before a major event, which suggests that the sudden absence of information has been encoded to signal a 'heads-up' to the players in anticipation of an important future event.

Conductor Study

The first study using the *Conductor's Jacket* system took place during several weeks in February 1998, with a professional conductor during rehearsals of a youth orchestra in Boston. During the few minutes before each rehearsal, the subject fitted the jacket on himself, the sensors were adjusted, and the entire system was tested. Then for the duration of the three-hour rehearsal, the system was used to record numerous files of physiological data timed with the external video camera. Notes were taken during the data acquisition trials, which were used to correlate and analyze the data and video files afterwards.

Initial Data

Initial results indicate several promising features in the data, including clear separation between the expressive use of both hands, context-dependent variations in the respiration signal, and enticing indicators of emotional arousal in skin conductance. Out of more than twenty-two files that were recorded, four have been analyzed in detail and found to contain useful correlations between expressive parameters and the musical score.

In general, the quality of the signals was surprisingly very good. The four EMG signals demonstrated a particularly high signal-to-noise ratio; that is, if there was no observable motion, then the signal was generally almost completely flat. This signal clarity suggests that signal-processing algorithms could be developed to yield good results for the automatic recognition of the above features. Such work remains to be done; however, we present below some preliminary findings extracted from the data by inspection.

Among many observations of the data that have been documented, several features were found to be particularly noteworthy. The following section demonstrates these features with graphical data taken from several rehearsal segments; they have been analyzed for their correlations with known features of traditional conducting technique. These features indicate that our subject:

- used his left hand to provide supplementary information and expression
- suddenly withdrew gestural information when he intended to signal the onset of a major event
- showed fundamental differences in the way he made information-carrying gestures vs. non-information carrying gestures
- modulated the force output of his muscles when generating a beat gesture in order to indicate the overall loudness or intensity of the music at that beat
- modulated his respiration to express the phrasing in the music

In our first two examples, EMG signals from the right and left biceps demonstrate how the left hand was used to provide extra information to supplement the information given by the right hand. In the first example, our subject chose to modulate the meter from 4 to 2. At the moment just before he intended for the meter to change, he reached out his left hand (which was until that moment at his side) and reinforced the new meter with both hands. Figure 1, shown below, demonstrates how the previous faster meter (where only the right hand was used) transitioned to a slower meter as the left hand entered:

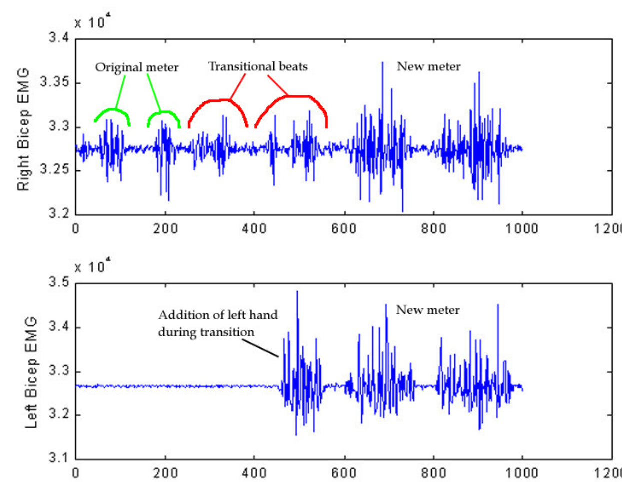

Figure 1. EMG signals from both biceps during a metrical shift.

The top graph shows the use of the right arm; in the first 200 samples of this segment, beats occur approximately every 100 samples. Then, during samples 220-600, the beats begin to transition to a new meter that is one-half as fast. These two beats are subdivided, as if to show both meters simultaneously. During the second of these beats, the left hand enters as if to emphasize the new tempo; this is shown in the bottom graph. Following this transition, the slower meter comes into relief (beginning at sample 600), with the new beat pattern showing a clearly defined envelope again.

In another section, our subject used his left hand to indicate a drastic reduction in loudness at the very end of the movement. As shown in Figure 2, below, the right hand gave all the beats leading up to the ending, but at the last minute the left hand was used to indicate a quick volume change and a quiet ending:

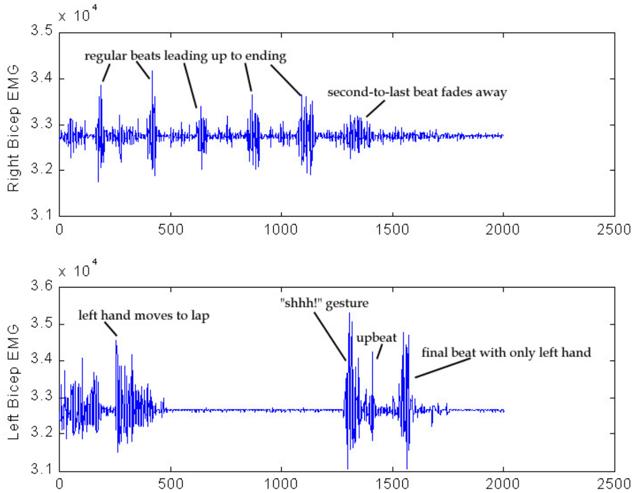

Figure 2. Use of the left hand to indicate drastic change in volume.

In this example, the right hand drops away at the very end and doesn't indicate the final beat. This drastic change in the use of the hands seems purposeful; the video shows that our subject looked directly at the wind section during this moment, as if he wanted to indicate a very different character for the final woodwind chords. As these first two examples have shown, the subject modified the *handedness* of his gestures in order to indicate something unusual.

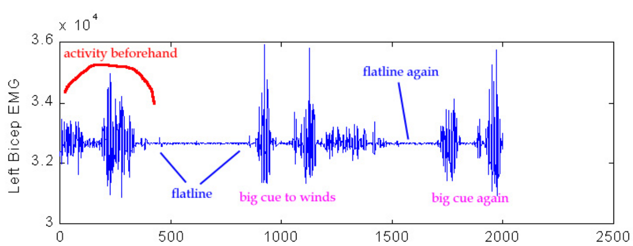

Figure 3. The characteristic "flatline" in the left bicep before a major event.

A second finding indicates that the subject would often withdraw gestural information suddenly before signaling the onset of a major event. That is, his EMG signals (particularly from the left bicep) often went to nearly zero before an important event or entrance. For example, it is very important for conductors to cue the woodwinds after they have waited silently for many measures; if the cue is not clear, they might not start playing in the right place. Such an event happens in bar number 32 of Prokofief's *Dance* movement; many of the winds need to play after many measures of silence. Leading up to this event, our subject used his left hand normally, and then, two measures before the wind entrance, stopped using it completely. Then, just in time for the cue, he gave a big pickup and downbeat with the left arm. In repetitions of the same passage, the same action is repeated. This is demonstrated in Figure 3.

A reasonable hypothesis for why this "flatline" occurs could be that the sudden lack of information is eye-catching for the musicians, and requires minimal effort from the conductor. The quick change between information-carrying and non-information-carrying states could be an efficient way of providing an extra cue ahead of time for the musicians.

A third feature we discovered in the EMG data is that the signals generated by the action of turning pages are inherently different in character from the signals generated by actions that are intended to convey musical information. That is, it seems as if page turns are done in such a way as to purposefully not attract attention or convey musical information. An example page turn is shown below in Figure 4:

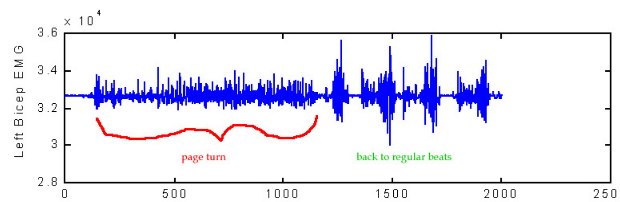

Figure 4. Difference between page turn gestures and information-carrying gestures.

We hope soon to be able to isolate aspects of these (and other non-useful) signals so as to teach a system how to distinguish an information-carrying from a non-information-carrying gesture.

A fourth feature found in the EMG signal is that the amplitude of a beat-generating spike seems to indicate intended sharpness of attack (or perhaps volume) of the notes to be played at that beat. This is compounded by what appears to be a kind of 'predictive' phenomenon, whereby the conductor indicates a very strong beat on the beat directly preceding the intended one. This is often discussed in the literature on conducting technique, but has never been shown quantitatively to be true. Figure 5, below, shows a segment of Prokofiev's *Dance* movement score with the accents highlighted and aligned with the accents given by our subject:

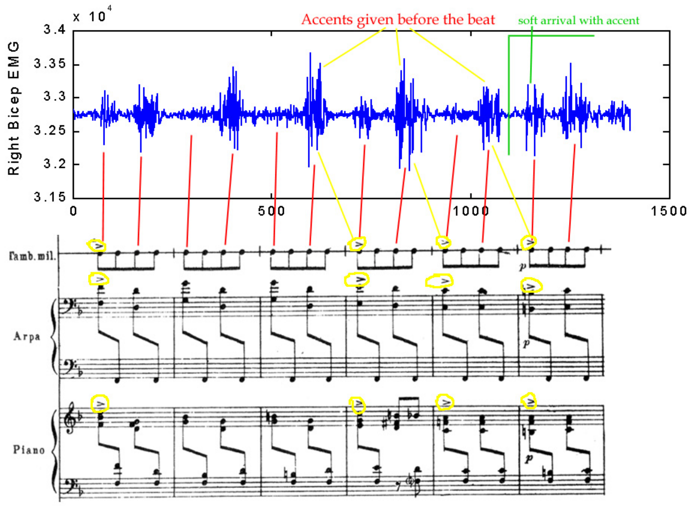

Figure 5. 'Predictive' accents and their relation to the score.

The dark, nearly-vertical lines show how the EMG signals line up with the beats in the score, while the lighter, slanted lines show the relationship between the conducted accent and the accent where it is to be played. The separation line around sample 1100 represents the barrier in-between the first, loud, section, and the second section, which is to be played quietly. The reduced amplitude of the EMG signal right before the separation line could indicate an anticipation of the new loudness level. Also, the existence of high-amplitude EMG signals on non-accented notes in this passage cannot be accounted for musically; perhaps they are due to conductor error. Alternately, they might be accounted for from the perspective of information theory, that once a pattern has been established it does not have to be indicated at every point.

Finally, we found correlations between the respiration signal and expressive aspects of the music. For example, in one musical section, our subject's respiration cycles matched the metrical cycles of the music; when the meter changed, so did his breathing patterns. Secondly, the amplitude of his respiration signal often increased in anticipation of a downbeat and sharply decreased right afterward. This might have been the result of the compression of the ribcage in the execution of the beat, but could also be an intentionally expressive phenomenon. For example, it is considered a standard practice among conductors to breathe in at upbeats and breathe out at downbeats, regulating their air flow relative to the speed and volume of the music.

Conclusions

The first results of the *Conductor's Jacket* project indicate several concrete findings. We are continuing to validate our preliminary results with additional data that has been taken from six other subjects who represent a range of abilities, techniques, and expressive styles.

In addition, a formal analysis of the data has been undertaken, which makes use of techniques from pattern recognition, signal processing (particularly short-time Fourier analysis), and semantic interpretation. Ultimately, we plan to synthesize models for musical performance and expression that incorporate affect and gesture, and that might be useful for the stage. Ideally, our results will be applicable both to professional conductors (enabling the composition of new works for conductors and orchestras where the conductor takes a more *instrumental* role) and to modern, technologically-augmented performers using physiological and gesture-capture systems.

Acknowledgements

We would like to thank the musicians who have participated in our studies, as well as Gregory Harman and Jennifer Healey for their technical assistance, and Professor Tod Machover for his advice. This work is supported in part by the "Things That Think" Consortium at the M.I.T. Media Lab.

References

Campbell, D. G. "Basal emotion patterns expressible in music." American Journal of Psychology 1942, 55, 1-17.

1. Hatfield E., J. Cacioppo, and R. Rapson. Emotional Contagion; Studies in Emotion and Social Interaction. Cambridge, UK: Cambridge University Press, 1994.
2. Healey, Jennifer and Rosalind Picard. "Digital Processing of Affective Signals." Appears in ICASSP '98, Seattle, WA.
3. Marrin, Teresa and Rosalind Picard. "The Conductor's Jacket: A Device For Recording Expressive Musical Gestures." Proceedings of the International Computer Music Conference, October 1998.
4. Picard, Rosalind. Affective Computing. Cambridge: M.I.T. Press, 1997.
5. Picard, Rosalind and J. Healey. "Affective Wearables." IEEE ISWC Proceedings vol. 1, no. 1, October 1997.