

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Sound Production for the Emotional Expression of Socially Interactive Robots

¹Eun-Sook Jee, ²Yong-Jeon Cheong, ³Chong Hui Kim,
²Dong-Soo Kwon, and ⁴Hisato Kobayashi

¹*The Future Robot Research Institute,*

²*Human-Robot Interaction Research Center, KAIST,*

³*Agency for Defense Development,*

⁴*Graduate School of Art and Technology, Hosei University,*

^{1,2,3}*Republic of Korea*

⁴*Japan*

1. Introduction

With the remarkable advancements in the field of robotics, the application of robots is no longer restricted to industrial automation but has been extended to personal home services. Robots are built to interact with humans, since they have not been developed to function as automatic machines, but to coexist as in human society. (Kim et al. 2005) Emotional interaction with humans is an integral function of socially interactive robots like Silbot—an intelligent robot developed in Korea for the purposes of assistance and entertainment geared toward the silver generation. When robots can comprehend human emotion and express their own emotion naturally, an emotional bond between human and robot is established.

Sounds and gestures are the two most basic mediums of emotional communication, and human beings constitute the subject of most studies on emotional communication. Many researches try to investigate the association of human emotion with a voice or facial expression. In recent years, the emotional aspects of music are being studied in both scientific and psychological contexts because of the complexity of emotional experiences in music. (Juslin & Sloboda 2001) In addition, a few researches focus on how to enable a robot to express emotion using speech synthesis, facial expressions, or sound. (Nakanishi & Kitagawa 2006; Jee et al. 2007)

The purpose of this section is to discuss not only the emotional sound design but also the process of emotional sound production aimed to enable robots to express emotion effectively, for facilitating the interaction between humans and robots. To begin with, we use the explicit or implicit link between emotional characteristics and musical parameters to compose six emotional sounds and then analyze them to identify a method to improve a robot's emotional expressiveness.

First, we introduced three emotional sounds—happiness, sadness, and fear—in robots, taking into consideration several musical parameters, namely, mode, tempo, pitch, rhythm, harmony, melody, volume, and timbre. Using the sound samples, we performed an

experiment to identify whether the sounds composed convey positive or negative emotions in the robot. Following this, we tested whether three basic emotional sounds coincided with the robot's facial expressions, using the Likert scaling method. This is another approach in the study of emotional expressiveness in robots. The results of experiments using either auditory or visual stimuli will then be compared with the results of experiments using both types of stimuli.

Second, we suggest the idea of incorporating intensity variation in emotional sounds with three different degrees: strong, middle, and weak. For this purpose, we produced additional emotional sounds of joy, shyness, and irritation. We regulate only three musical parameters—tempo, pitch, and volume—because of the technical limitations of the computer system; in other words, robots can control only the tempo, pitch, and volume. Although only three parameters can be regulated and manipulated in our set up, the intensity variation causes more dynamic emotional human-robot interaction.

Finally, we present the idea of synchronization with the emotional sounds of joy, shyness, and irritation. The synchronization of emotional sounds with the behavior of robots, including their movements and gestures, is a key issue in real implementation because it makes a robot's behavior more natural. For the synchronization, we divided the emotional sounds into several segments in accordance with musical structure. Some segments of the emotional sounds are repeatable and robots can control their sound duration for the synchronization.

2. Previous work

Can you ever imagine a movie without sound? Sound is an integral element of human communication and interaction. From the perspective of cognitive science, among human activities related to sound, both language and music share many things in common. For instance, both language and music unfold sound in time. Similar to language, music has a hierarchical structure and it is also believed to have a grammatical structure. (Lerdahl & Jackendoff 1983) What then is the real difference between music and language? Perhaps, the most conspicuous difference is that music has an emotional meaning and induces genuine and deep emotions. (Meyer 1956) There is no language more powerful than the language of music. Music is indeed a language of emotion. (Pratt 1948)

As an emotionally rich medium, music is necessary to express a robot's emotion for human-robot interaction. There exist numerous studies on music and emotion since time immemorial because this topic has always interested people. There are multidisciplinary approaches to understand music and emotion because the emotional experience of music is complex and rich. Several scholars have developed aesthetic and philosophical discussions on music and emotion. (Davies 2001; Kivy 1999; Levinson 1982) Musicologists and music theorists have studied emotional expressiveness not only in western art music but also in pop music. (Cook & Dibben 2001; Meyer 1956) Further, Feld (1982) and Becker (2001) have approached music and emotion through anthropological and ethno-musicological perspectives. DeNora (2001) has applied sociology paradigms to understand the relation between emotion and music. Finally, Bunt & Pavlicevic (2001) have studied music and emotion for therapeutic purposes.

Psychological perspectives on music and emotion can be examined in further detail because emotion is a main concern of psychology. Recently, large-scale investigations on the relation between music and emotion have been performed through psycho-biological or neuro-

psychological approaches. For instance, using the technique of positron emission tomography (PET), Blood et al. (1999) examined the change in cerebral blood flow during emotional responses to music. They found that music could recruit the neural mechanisms associated with pleasant or unpleasant emotional states. Baumgartner et al. (2006) investigated how music enhances emotions using functional magnetic resonance imaging (fMRI). The brain imaging showed that visual and musical stimuli automatically evoke strong emotional feelings and experiences. Peretz (2006) presented the neural correlates of musical emotion and the existence of specific neural arrangements for certain emotions induced by music. Juslin & Västfall (2008) determined the existence of underlying mechanisms in music that evoke emotions and concluded that these mechanisms are not unique to music. Livingston & Thomson (2009) suggested that music generates emotional experiences by activating the channels related to the audio-visual neuron system. In addition, some psychologists have studied which musical parameters evoke emotional feelings. For instance, Hevner (1935; 1936; 1937) researched the emotional meanings in music through psychological experiments. She created what she termed the Adjective Circle by categorizing of emotions into eight adjective groups, as shown in Fig. 1.



Fig. 1. Hevner’s categorization of emotions: the adjective circle

Hevner assumed some associations between musical parameters and emotion. Through experiments, she found that a specific musical parameter was responsible for a particular emotional response. Hevner considered six musical parameters—mode, tempo, pitch, rhythm, harmony, and melody; we have carefully analyzed these during emotional sound production. These parameters and their associations with emotion are briefly summarized as follows.

- Mode is any of the certain fixed arrangements of tones, such as major or minor. Major modes manifest gracefulness (c5) and happiness (c6), while minor modes indicate sadness (c2) and sentimentality (c3).
- Tempo is the speed of music. Fast tempi signify happiness (c6) and excitement (c7), whereas, slow tempi indicate solemnity (c1), sadness (c2), sentimentality (c3), and serenity (c4).
- Pitch is the frequency of sound. Pitches in higher register express serenity (c4) and gracefulness (c5), whereas, pitches in lower register represent sadness (c2) and vigorousness (c8).
- Rhythm is the aspect of music that comprises all the elements that relate to forward movement. A firm rhythm indicates solemnity (c1) and vigorousness (c8). On the contrary, a flowing rhythm expresses sentimentality (c3), gracefulness (c5), and happiness (c6).
- Harmony is the combination of simultaneous musical notes in a chord. A simple harmony represents serenity (c4), gracefulness (c5), and happiness (c6), whereas a complex harmony expresses sadness (c2), excitement (c7), and vigorousness (c8).
- Melody is a succession of single notes that form a tune. Ascending melodies signify solemnity (c1) and serenity (c4), whereas, descending melodies express gracefulness (c5), excitement (c7), and vigorousness (c8).

Hevner's pioneering experimental researches on music and emotion continue to intellectually stimulate researchers today. Interestingly, Juslin (2000) studied the utilization of acoustic cues in the communication of musical emotions between performer and listener and measured the correlation between various emotional expressions and acoustic cues. Gabrielsson & Lindström (2001) presented a historical overview of studies on musical structures and emotion. They suggested more specific musical parameters than Hevner. For example, Gabrielsson & Lindström examined tempo, mode, loudness, pitch, intervals, melody, harmony, tonality, rhythm, timbre, articulation, amplitude envelope, musical form, and the interaction between parameters. They summarized the relation between newly arranged musical parameters and emotion. Juslin & Laukka (2003) modeled the emotional expression of different music performances by means of multiple regression analysis, to clarify the relationship between emotional descriptions and measured parameters such as tempo, sound level, and articulation. Similarly, Schubert (2004) considered different musical parameters of loudness, tempo, melodic contour, texture, and timbre. He investigated the relationship between these parameters and perceived emotion by using continuous response methodology and time-series analysis. In the area of computer entertainment, Berg & Wingstedt (2005) mentioned that the influence of visual aspects on emotional dimensions has been researched more systematically than that of sound. They simulated the relation between musical parameters and expressed emotions by selecting mode, instrumentation, tempo, articulation, volume, and pitch register as musical parameters. Further, through an examination of over 100 empirical studies, Livingstone & Thompson (2006) concluded the corresponding associations between specific emotions and musical parameters. Their results are similar to that of Gabrielsson & Lindström. In addition, Post & Huron (2009) found that the minor mode in Western classical music tends to be slower, based on Hevner's theory and Juslin's cue utilization that the minor mode is associated with sadness (c2) and sentimentality (c3).

As we examined above, the study of emotion and music has a short history. Studies on emotional expression through musical sounds in robotics are even rarer. In the following

three sections, we will specify our processes of emotional sound productions in order to enhance a robot’s expressiveness through emotional sound coincidence with facial expression, intensity variation of emotional sounds, and sound synchronization with the robot’s behavior.

3. Production of basic emotional sounds

In this section, we present the design and production of a robot’s emotional sounds. The duration of each emotional sound is two or three seconds. Emotional sounds are produced by MIDI, sound filtering or mixing. The raw audio samples are recorded and filtered through Sound Forge and Cubase software. Some of the filtered audio samples are then mixed with pre-recorded midi sounds by Cubase in order to create emotional sounds in a robot. Fig. 2 shows the process of the emotional sound production.

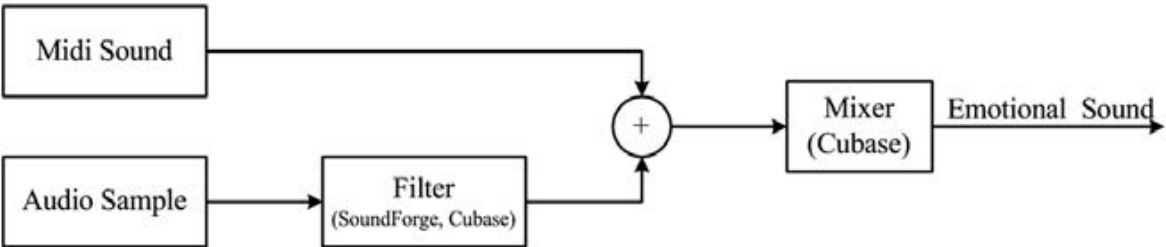


Fig. 2. The design flow of sound production

Basic emotions are defined as a limited number of innate and universal categories or emotions from which all other emotional states can be derived. (Cited in Berg & Wingstedt 2005) Juslin & Sloboda (2001) discussed that basic emotions belong to at least five categories: happiness, sadness, anger, fear, and disgust. We decided to produce two sets of three

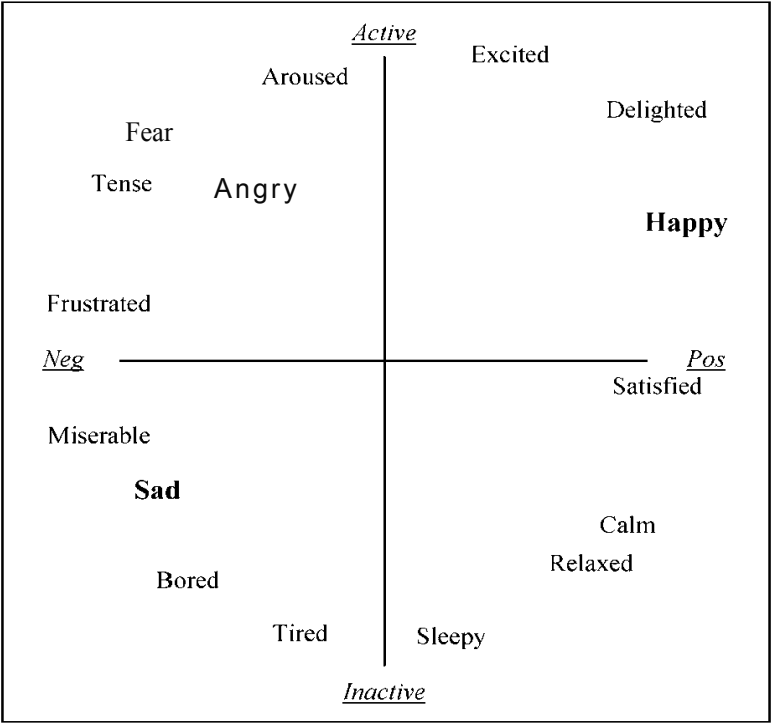


Fig. 3. Two dimensional circumplex model of emotions

emotional sounds: (1) happiness, sadness, and fear and (2) joy, shyness, and irritation. The first set is produced to test the effect of emotional sounds and how these sounds coincide with facial expressions. The second set pertains to the intensity variation of emotional sounds, and the synchronization of the sounds with a robot's behavior. Each emotional sound in both groups is located on three different sections of a two-dimensional circumplex model of emotion, involved in the dimensions of arousal (activity) and valence (positive/negative). (Russell 1980) Figure 4 presents the two-dimensional circumplex model of emotion. With respect to this model, happiness of set 1 and joy of set 2 represent an active and positively valenced emotion, while sadness of set 1 and shyness of set 2 denote an inactive and negatively valenced emotion. Happiness and joy are symmetrically opposite to sadness and shyness. Besides them, we also decided to produce emotional sounds for fear of set 1 and irritation of set 2, which are opposite to happiness and joy on the valence perspective and also opposite to sadness and shyness on the arousal perspective.

On the basis of prior investigations on which musical sound evokes emotion, the following musical parameters will be examined for the three basic emotional sounds of set 1: Hevner's six musical parameters, volume, and timbre. As mentioned above, Hevner considered mode (major or minor), tempo (fast or slow), pitch (high or low), rhythm (firm or flowing), harmony (simple or complex), and melody (ascending or descending). In addition, we examine volume and timbre. As an aside, note that timbre can be defined as an instrumental setting.

3.1 Happiness

The sound of our happiness is in the quasi-major mode. The tempo, at 160 BPM ($J = 160$), is very fast owing to the subdivisions of quarter note (i.e., eighth notes, triplets, and sixteenth notes). Most of the notes are in the high pitch range from E4 (ca. 329.6 Hz) to F#6 (ca. 1174.6 Hz). The harmony is simple with major triads, and the rhythm is firm with a vibraphone's quarter notes on beat. Happiness has an ascending melodic contour, and the volume of happiness is 60 dB SPL (10^{-6} watt/m²). Sounds from the ocarina and vibraphone, produced using a midi keyboard, are used for the timbre of happiness. Fig. 4 shows the score of the sound of happiness.



Fig. 4. Music score for happiness

3.2 Sadness

The sound of sadness is neither in the major nor minor mode. The tempo is 99 BPM ($J = 99$) and very slow because sadness consists of 1 quarter note and 2 dotted half notes. The pitch ranges from G4 (ca. 155.6 Hz) to C7 (ca. 1046.5 Hz). The harmony is complex because of the absence of major or minor triads, and the rhythm is firm with 2 downbeat dotted half notes. The melody of sadness is descending, and the volume of sadness is the same as that of happiness as 60 dB SPL (10^{-6} watt/m²). The cello and piano are used to determine the timbre. Fig. 5 shows the score of a sadness sound.



Fig. 5. Music score for sadness

3.3 Fear

Similar to sadness, the sound of fear is neither in major nor minor mode because we intend to express the negative valence of sadness and fear by using the same melody line. The tempo is 126 BPM (♩ = 126) but, in reality, it is slower than the tempo of sadness because fear consists of only dotted half notes. Moreover, the duration of the last note is tripled by a tie. The pitch is the lowest among the emotional sounds that we produce. It ranges from G2 (ca. 97.9 Hz) to A3 (ca. 233.1 Hz). The harmony of fear is very simple because only octaves (1:1 ratio) are used. The rhythm is very firm with only downbeat notes, and the melody of fear is descending. The volume of happiness is 70 dB SPL (10⁻⁵ watt/m²). In this case, the organ is used to determine the timbre. In the last long note, the vibration that is characteristic of an organ timbre, is fully revealed. Fig. 6 shows the score of the fear sound.



Fig. 6. Music score for fear

3.4 Experiment on basic emotional sounds

We conducted an experiment to test whether emotional sounds evoke or induce happiness, sadness, and fear. We recruited 20 participants, comprising an equal number of men and women. Our participants were asked to rate their emotional states on the Likert five-point scale after listening to randomly presented sounds. The experiment revealed that 90% made positive responses on our happiness sound; more than half of the participants rated this sound very strongly. On our sadness sound, 65% reported a strong feeling of sadness. Further, 50% of the participants responded positively to the sound of fear, and among them, 15% rated the sound very strongly. Table 1 shows how effectively the sounds express the three basic emotions, from the results of the experiment.

	Happiness	Sadness	Fear
Never			
Weak		2 (10%)	3 (15%)
Moderate	2 (10%)	5 (25%)	7 (35%)
Strong	7 (35%)	13 (65%)	7 (35%)
Very Strong	11 (55%)		3 (15%)
Sum	20 (100%)		

Table 1. Sound validity of the three basic emotions

3.5 Experiment on the coincidence of basic emotional sounds with facial expressions

Nakanishi et al. (2006) proposed a visualization of musical impressions on faces in order to represent emotions. They developed a media-lexicon transformation operator of musical data to extract some impression words from musical elements that determine the form or structure of a song. Lim et al. (2007) suggested the emergent emotion model and described some flexible approaches to determine the generation of emotion and facial mapping. They mapped the three facial features of the mouth, eyes, and eyebrows into the arousal and valence of the two-dimensional circumplex model of emotions.

Even if robots express their emotions through facial expressions, their users or partners could face a problem perceiving the subtle differences in a given emotion. The subtle change of emotion is difficult to perceive through facial expressions, and hence, we selected several representative facial expressions that people can understand easily. Coinciding basic emotional sounds with the facial expression of robots is, hence, an important issue. We performed the experiment to test the whether the basic emotional sounds of happiness, sadness, and fear coincide with the corresponding facial expressions.

We then compared the results of the experiment against either basic emotional sounds or facial expressions with both sounds and facial expression. The experiment on the coincidence of sounds and facial expressions was performed on the same 20 participants. Since the entire robot system is still in its developmental stage, we conducted the experiments using laptops, on which we displayed the facial expressions of happiness, sadness, and fear, following which we played the music composed as part of the preliminary experiment. Figure 8 shows the three facial expressions we employed for the experiment.

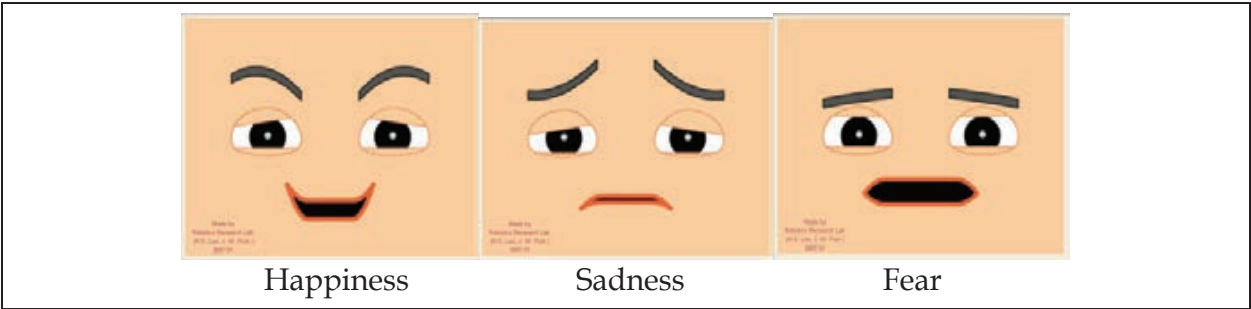


Fig. 7. Facial expressions of a preliminary robot

Table 2 shows the results on the coincidence of musical sounds and the facial expressions of happiness, sadness, and fear. The results supported our hypothesis on the coincidence of basic emotional sounds with facial expressions. For instance, a simultaneous simulation of sound and the facial expression of fear show a more positive improvement than that of either sound or facial expression. Therefore, the sounds and facial expressions cooperate complementarily for the conveyance of emotion.

4. Intensity variation of emotional sounds

Human beings are not keenly sensitive to detecting the gradual change in sensory stimuli that evoke emotions. Delivery of delicate changes in emotions through both facial expressions and sounds is difficult. When comparing the conveying of delicate emotional changes, sound is more effective than facial expressions. Cardoso et al. (2001) measured the intensity of emotion through experiments using numerical magnitude estimation (NE) and

	Sound			Facial Expression			Sound with Facial Expression		
	Happiness	Sadness	Fear	Happiness	Sadness	Fear	Happiness	Sadness	Fear
Never						2 (10%)			
Weak		2 (10%)	3 (15%)	1 (5%)	1 (5%)	4 (20%)			
Moderate	2 (10%)	5 (25%)	7 (35%)	7 (35%)	5 (25%)	6 (30%)	4 (20%)	3 (15%)	2 (10%)
Strong	7 (35%)	13 (65%)	7 (35%)	12 (60%)	12 (60%)	8 (40%)	8 (40%)	11 (55%)	10 (50%)
Very Strong	11 (55%)		3 (15%)		2 (10%)		8 (40%)	6 (30%)	8 (40%)
Sum	20 (100%)								

Table 2. Coincidence of emotional sounds and facial expressions

cross-modal matching to line-length responses (LLR) in a more psychophysical approach. We quantized the levels of emotional sounds as strong, middle, and weak, or strong and weak in terms of intensity variation. The intensity variation is regulated on the basis of the result of Kendall’s coefficient between NE and LLR. (Cardoso et al. 2001) Through the intensity variation of the emotional sounds, robots can express delicate changes in their emotional state.

We already discussed several different musical parameters for sound production and for displaying a robot’s basic emotional state in section 3. Among these, only three musical parameters—tempo, pitch, and volume—are related to intensity variation because of the technical limitations of the robot’s computer system. Our approach to the intensity variation of the robot’s emotions is introduced with the three sound samples of joy, shyness, and irritation, which are equivalent to happiness, sadness, and fear on the two-dimensional circumplex model of emotion.

First, volume was controlled in the range from 80~85% to 120~130%. When the volume of any sound is changed beyond this range, the unique characteristic of emotional sound is distorted and confused.

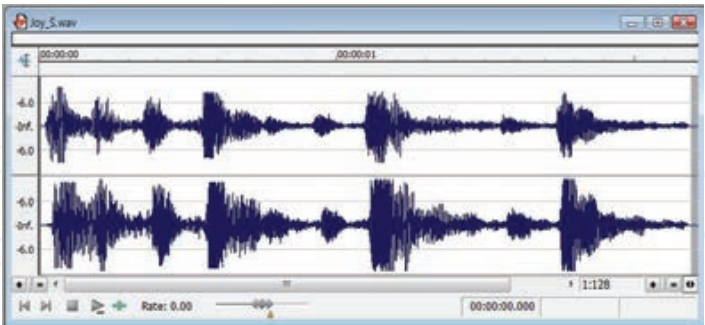
Second, in the same way as volume regulation, we controlled the tempo to within the range of 80~85% to 120~130% of middle emotional sounds. When the tempo of the sound changes to slower than 80% of the original sound, the characteristic of the emotional state of the sound disappears. Reversely, when the tempo of the sound accelerates and is faster than 130% of the original sound, the atmosphere of the original sound is modified.

Third, the pitch was also controlled but the change of tempo and volume is more distinct and effective for intensity variation. We only changed the pitch of irritation because the sound of irritation is not based on the major or minor mode. The sound cluster in the irritation sound moves with a slight change in pitch in glissando.

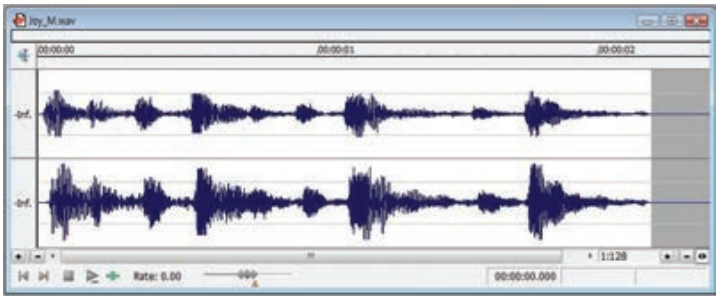
4.1 Joy

Joy shares common musical characteristics with happiness. For the middle joy sound, the mode is the quasi major. The tempo is 116 BPM (J = 116) and is quite fast in real life because of the triplets. The pitch ranges from D3 (ca. 146.8 Hz) to C5 (ca. 523.3 Hz). The rhythm is firm with on-beat quarter notes. The harmony is simple owing to major triads, the melody is

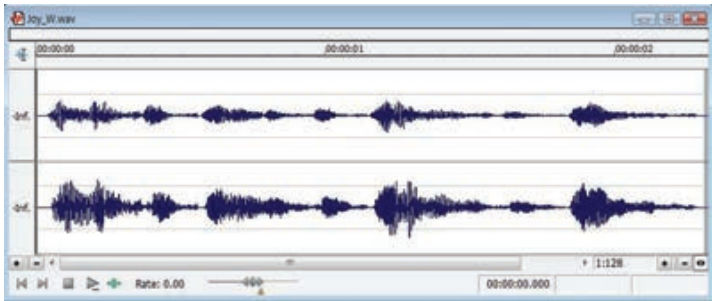
ascending, and the volume is 60 dB SPL (10^{-6} watt/m²). The staccato and pizzicato of string instruments determine the timbre of the sound of joy. Figure 8 illustrates wave files depicting strong, middle, and weak levels of joy.



(a) Strong Joy



(b) Middle Joy



(c) Weak Joy

Fig. 8. Wave file depicting strong, middle, and weak joy sound samples

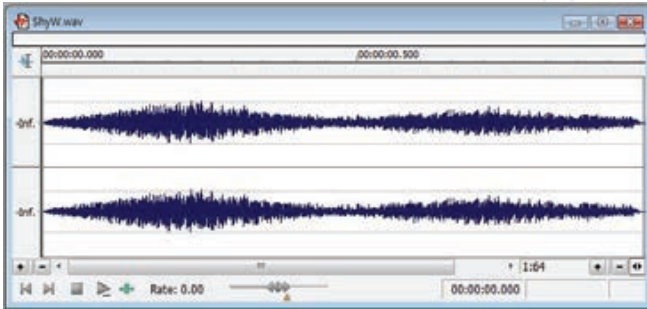
For the emotion of strong joy, the volume is only increased to 70 dB SPL (10^{-6} watt/m²). On the other hand, for a weak joy emotion, we decrease the volume down to 50 dB SPL (10^{-7} watt/m²) and reduce the tempo. Table 3 shows the change in the musical parameters of tempo, pitch, and volume for intensity variation of the sound for joy.

Intensity	STRONG	Middle	Weak
Volume	120% 70 dB SPL	100% 60 dB SPL	80% 50 dB SPL
Tempo	100%	100%	120%
Pitch	146.8~523.3Hz		

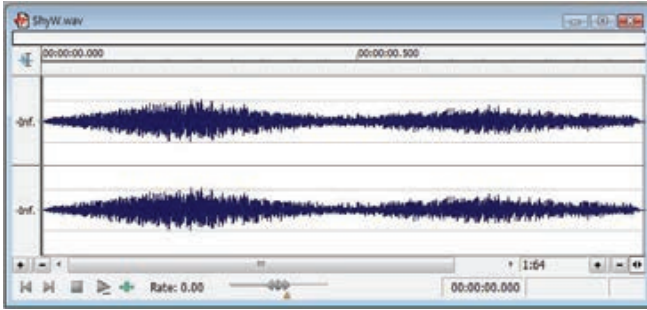
Table 3. Intensity variation of joy

4.2 Shyness

Shyness possesses emotional qualities similar to sadness on the two-dimensional circumplex model of emotion. The intensity variation of shyness is performed on two levels: strong and weak. As a standard, a strong shyness sound is composed on the basis of neither a major nor minor mode because a female voice is recorded and filtered in this case. The tempo is 132 BPM ($\text{♩} = 132$). The pitch ranges from Bb4 (ca. 233.1 Hz) to quasi B5 (ca. 493.9 Hz). The rhythm is firm, the harmony is complex with a sound cluster, and the melody is a descending glissando with an obscure ending pitch point. The volume is 60 dB SPL (10^{-6} watt/m²) and the metallic timbre is acquired through filtering. Figure 9 shows the wave files of strong shyness and weak shyness.



(a) Strong Shyness



(b) Weak Shyness

Fig. 9. Wave file depicting strong and weak shyness sound samples

For weak shyness, the volume is reduced to 50 dB SPL (10^{-7} watt/m²), and the tempo is also reduced. Table 4 shows the intensity variation of shyness.

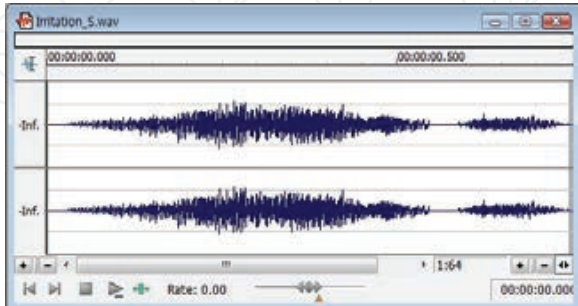
Intensity	STRONG	Weak
Volume	100%	80%
Tempo	100% 60 dB SPL	115% 50 dB SPL
Pitch (Semitone)	233.1~.493.9Hz	

Table 4. Intensity variation of shyness

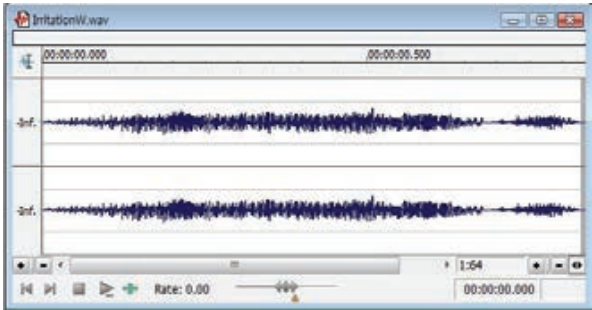
4.3 Irritation

The emotional qualities of irritation are similar to those of fear. Irritation also only has two kinds of intensity levels. Strong irritation, as a standard sound, is composed on the basis of

neither the major nor minor mode because it constitutes a combined audio file and midi featuring a filtered human voice. The tempo is 112 BPM ($J = 112$), and the pitch ranges from C4 (ca. 261.6 Hz) to B5 (ca. 493.9 Hz). The rhythm is firm, and the harmony is complex with a sound cluster. The melody is an ascending glissando, which is the opposite of shyness. It reflects an opposite status on the arousal dimension. The volume is 70 dB SPL (10^{-5} watt/m²), and the metallic timbre is acquired through filtering, while the chic quality of timbre comes from a midi. Figure 10 shows wave files of strong and weak irritation.



(a) Strong Irritation



(b) Weak Irritation

Fig. 10. Wave files depicting strong and weak irritation sound samples

For the weak irritation sample, the volume is decreased to 60 dB SPL (10^{-6} watt/m²) and the tempo is reduced. Table 5 shows how we regulated the intensity variation of irritation.

Intensity	STRONG	Weak
Volume	100% 70 dB SPL	85% 60 dB SPL
Tempo	100%	115%
Pitch	261.6~493.9 Hz	220~415.3 Hz

Table 5. Intensity variation of irritation

5. Musical structure of emotional sounds to be synchronized with a robot’s behavior

The synchronization of the duration of sound with a robot’s behavior is important to ensure the natural expression of emotion. Friberg (2004) suggested a system that could be used for analyzing the emotional expressions of both music and body motion. The analysis was done in three steps comprising cue analysis, calibration, and fuzzy mapping. The fuzzy mapper translates the cue values into three emotional outputs: happiness, sadness, and anger.

A robot’s behavior, which is important in depicting emotion, is essentially continuous. Hence, for emotional communication, the duration of emotional sounds should be synchronized with that of a robot’s behavior including motions and gestures. At the beginning of sound production, we assumed that robots could control the duration of their emotional sounds. On the basis of the musical structure of sound, we intentionally composed the sound such that it consists of several segments. For the synchronization, the emotional sounds of joy, shyness, and irritation have musically structural segments, which can be repeated as per a robot’s volition. The most important considerations for synchronization are as follows:

1. The melody of emotional sounds should not leap abruptly.
2. The sound density should not be changed excessively.
 - If these two points are not retained, the separation of the segment would be difficult.
3. Each segment of any emotional sound contains a specific musical parameter which is peculiar to the quality of the emotion.
4. Among the segments of any emotional sound, the best segment containing the characteristic quality of the emotion should be repeated.
5. When a robot stretches a sound by repeating one of the segments, both the repetition and the connection points should be connected seamlessly without any clashes or noises.

5.1 Joy

We explain our approach to synchronization by using the three examples of joy, irritation, and shyness, which are presented in section 4. As mentioned above, each emotional sound consists of segments that are in accordance with the musical structure. The duration of the joy sound is about 2.07s, and joy is divided into three segments: A, B, and C. Robots could regulate the duration of joy by calculating the duration of their behavior and repeating any segment to synchronize it. The figure of segment A is characterized by ascending triplets, and its duration is approximately 1.03s. Segment B is denoted by the dotted notes, and the duration of both segments B and C is about 0.52s. Figure 11 shows the musical structure of joy and its duration.



Fig. 11. Musical segments and the duration of joy

5.2 Shyness

The duration of shyness is about 1s. Shyness has two segments, A and B. The figure of segment A is characterized by a descending glissando on the upper layer and a sound

cluster on the lower layer. Segment B only has a descending glissando without a sound cluster on the lower layer. The duration of both segments A and B is about 0.52s. Figure 12 shows the musical structure of shyness and its duration.

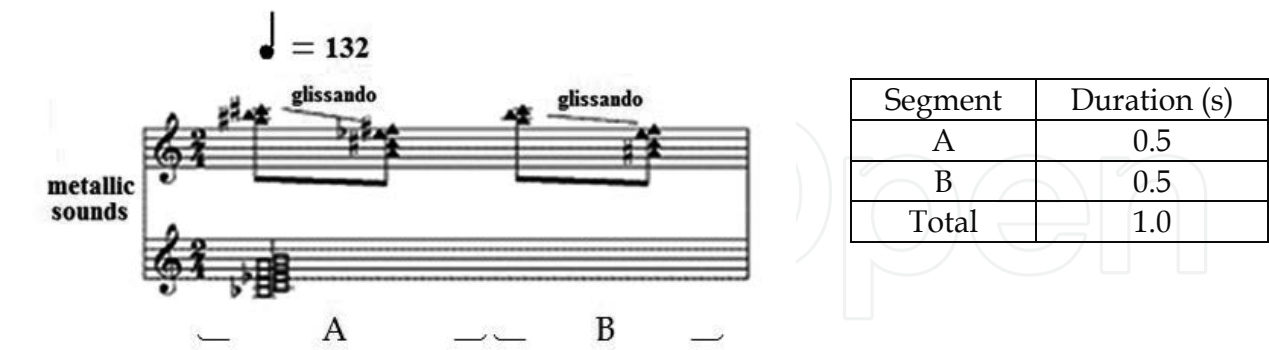


Fig. 12. Musical segments and the duration of shyness

5.3 Irritation

Irritation has almost the same structure as that of shyness. The duration of irritation is about 1.08s. Irritation has two segments, A and B. The figure of segment A is characterized by an ascending glissando. Segment B has one shouting. The duration of both segments A and B is about 0.54s. Figure 13 shows the musical structure of shyness and its duration.

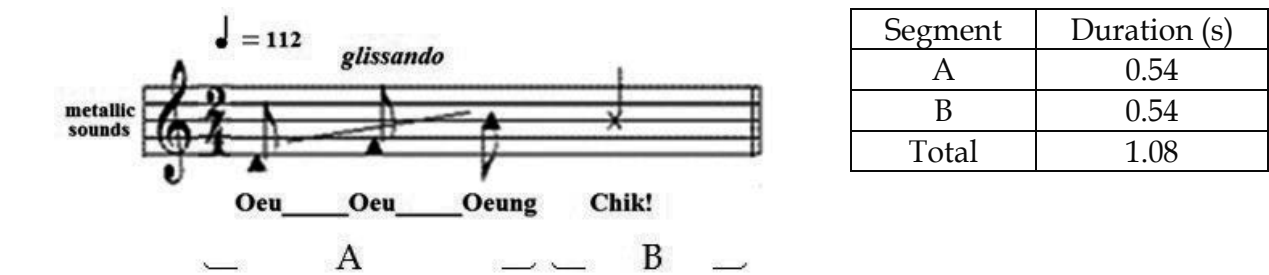


Fig. 13. Musical segments and the duration of irritation

6. Conclusion

In conclusion, the paper presents three processes of sound production needed to enable emotional expression in robots. First, we consider the relation between three basic emotions of happiness, sadness, and fear, and eight musical parameters of mode, tempo, pitch, rhythm, harmony, melody, volume, and timbre. The survey using the 5-point Likert scale, which was administered to 20 participants, proved the validity of Silbot’s emotional sound. In addition, the synchronizing of the robot’s basic emotional sounds of happiness, sadness, and fear with facial expressions is tested through the experiment. The results support the hypothesis that the simultaneous presentation of sound samples and facial expressions is more effective than the presentation of either sound or facial expression. Second, we produced emotional sounds for joy, shyness, and irritation in order to determine the intensity variation of the robot’s emotional state. Owing to the technical limitations of the computer systems controlling the robot, only three musical parameters of volume, tempo, and pitch are regulated for intensity variation. Third, the synchronization of the durations of

sounds depicting joy, shyness, and irritation with the robot's behavior is obtained to ensure a more natural and dynamic emotional interaction between people and robots.

7. References

- Baumgartner, T.; Lutz, K.; Schmidt, C. F. & Jäncke, L. (2006). The emotional power of music: How music enhances the feeling of affective pictures, *Brain Research*, Vol. 1075, pp. 151-164, 0006-8993
- Berg, J. & Wingstedt, J. (2005). Relations between selected musical parameters and expressed emotions extending the potential of computer entertainment, In the Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, pp. 164-171
- Blood, A. J.; Zatorre, R. J.; Bermudez, P. & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions, *Nature Neuroscience*, Vol. 2, No. 4, (April) pp. 382-387, 1097-6256
- Cardoso, F. M. S.; Matsushima, E. H.; Kamizaki, R.; Oliveira, A. N. & Da Silva, J. A. (2001). The measurement of emotion intensity: A psychophysical approach, In the Proceedings of the Seventeenth Annual Meeting of the International Society for Psychophysics, pp. 332-337
- Feld, S. (1982). *Sound and sentiment: Birds, weeping, poetics, and song in Kaluli expression*, University of Pennsylvania Press, 0-8122-1299-1, Philadelphia
- Hevner, K. (1935). Expression in music: A discussion of experimental studies and theories, *Psychological Review*, Vol. 42, pp. 186-204, 0033-295X
- Hevner, K. (1935). The affective character of the major and minor modes in music, *American Journal of Psychology*, Vol. 47, No. 4, pp. 103-118, 0002-9556
- Hevner, K. (1936). Experimental studies of the elements of expression in music, *American Journal of Psychology*, Vol. 48, No. 2, pp. 248-268, 0002-9556
- Hevner, K. (1937). The affective value of pitch and tempo in music, *American Journal of Psychology*, Vol. 49, No. 4, pp. 621-630, 0002-9556
- Jee, E. S.; Kim, C. H.; Park, S. Y. & Lee, K. W. (2007). Composition of musical sound expressing an emotion of robot based on musical factors, *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication*, pp. 637-641, ISBN, Jeju, Aug. 2007, Republic of Korea
- Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: relating performance to perception, *Journal of Experimental Psychology*, Vol. 16, No. 6, pp. 1797-1813, 0096-1523
- Juslin, P. N. & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, Vol. 129, No. 5, pp. 770-814, 0033-2909
- Juslin, P. N. & Sloboda, J. A. (Ed.) (2001). *Music and emotion*, Oxford University Press, 978-0-19-2263189-3, Oxford
- Juslin, P. N. & Västfall, D. (2008). Emotional responses to music: The need to consider underlying mechanisms, *Behavioral and Brain Sciences*, Vol. 31, pp. 556-621, 0140-525X
- Kim, H. R.; Lee, K. W. & Kwon, D. S. (2005). Emotional interaction model for a service robot, *Proceedings of the IEEE International Workshop on Robots and Human Interactive Communication*, pp. 672-678, Nashville, United States of America

- Kivy, P. (1999). Feeling the musical emotions, *British Journal of Aesthetics*, Vol. 39, pp. 1-13, 0007-0904
- Lerdahl, F. & Jackendoff, R. (1983). *A generative theory of tonal music*, MIT Press, 026262107X, Cambridge, Mass.
- Levinson, J. (1982). Music and negative emotion, *Pacific Philosophical Quarterly*, Vol. 63, pp. 327-346, 0279-0750
- Livingstone, S. R.; Muhlberger, R.; Brown, A. R. & Loch, A. (2007). Controlling musical emotionality: An affective computational architecture for influencing musical emotions, *Digital Creativity*, 18, pp. 43-54
- Livingstone, S. R. & Thompson, W. F. (2009). The emergence of music from the theory of mind, *Musicae Scientiae Special Issue on Music and Evolution* in press. 1029-8649
- Meyer, L. B. (1956). *Emotion and meaning in music*. University of Chicago Press, 0-226-52139-7, Chicago
- Nakanishi, T. & Kitagawa T. (2006). Visualization of music impression in facial expression to represent emotion, *Proceedings of Asia-Pacific Conference on Conceptual Modelling*, pp. 55-64
- Post, O. & Huron, D. (2009). Western classical music in the minor mode is slower (except in the romantic period), *Empirical Musicology Review*, Vol. 4, No. 1, pp. 2-10, 1559-5749
- Pratt, C. C. (1948). Music as a language of emotion, *Bulletin of the American Musicological Society*, No. 11/12/13 (September, 1948), pp. 67-68, 1544-4708
- Russel, J. A. (1980). A circumplex model of affect, *Journal of Personality and Social Psychology*, 39, 1161-1178
- Schubert, E. (2004). Modeling perceived emotion with continuous musical features, *Music Perception*, Vol. 21, No. 4, pp. 561-85, 0730-7829
- Miranda, E. R. & Drouet, E. (2006). Evolution of musical lexicons by singing robots, *Proceedings of TAROS 2006 Conference - Towards Autonomous Robotics Systems*, Gilford, United Kingdom

© 2009 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen