Subjective Timbre Space of 45 Modified Violin Tones

*Chul-Yong Ahn, *Hee-Suk Pang, *Koeng-Mo Sung

Abstract

In this paper we studied the reduced subjective timbre space of time-varying tones as well as steady state tones. 45 modified test tones were constructed from the original violin tone in consideration of 4 physical factors: spectrum enveleope, inharmonicity, time-varying spectrum and time reversal. The semantic differential (SD) method was used in the listening test. According to the factor analysis, the adjectives can be factorized into 4 groups. The first factor is characterized by the adjectives, 'free', 'broad', 'deep', 'rich', 'strong' and 'reverberant', the second by 'tenor', 'clear', 'bright', 'light' and 'sharp', the third by 'easy', 'smooth', and 'solid', the fourth by 'warm' and 'full'. The first factor, 'richness', seems to be dependent upon the time-varying characteristic of a tone. The second factor, 'sharpness', is shown to play an important role in a time-varying tone as well as in a steady state tone.

Ⅰ. INTRODUCTION

Timbre is the attribute of auditory sensation in terms of which a listener can judge two sounds similarly presented and having the same loudness and pitch as dissimilar. Researchers have been interested in finding the relationship between the physical and psychological factors of a tone, but the difficulty lies in that both factors of a tone have the multidimensional characteristics. That is, timbre of a tone depends on the spectrum, waveform, relative ratios of harmonics, formant, envelope, dynamic spectrum and so on, which implies that the physical factors of a tone are multidimensional [1]. And we describe the timbre of a tone by adjectives such as sharp, warm, smooth and so on, which implies that the psychological factors of a tone are also multidimensional.

In timbre research, semantic differential (SD) and multi-dimensional scaling (MDS) methods have been frequently used. In the SD method, sounds are rated on category scales that are composed of opposite adjectives such as sharp-dull, bright-dark. Using this method and principal components analysis(PCA) , Bismarck found 'sharpness' of a steady-state tone which is one of the most important psychological factors of timbre [2,3]. On the contrary, MDS method is based on similarity ratings of sounds and one can analyze the results by investigating the groups of common attribute tones. Grey found that timbre of a tone was interpretable by the spectral energy distribution, the presence of synchronicity in the transients of the harmonics and spectral fluctuation through time, and the presence of low-amplitude, high frequency energy in attack [1,4]. Iverson et al. found that both attack and steady state of a tone contribute to the perceptual dimension [5].

The results of the above researches have not considered the time-varying tones with the SD method and then, this paper investigated the subjective timbre space of time-varying tones as well as steady-state tones by using SD method and factor analysis. First, we make 45 test tones by modifying an original violin tone and then perform the listening tests. Second, we analyze the listening test results by factor analysis and discuss the 4 factors obtained. Finally we make conclusion.

Ⅱ. Experiment

2.1. Subjects

20 male subjects majoring in acoustics with normal hearing participated in the listening test. About half of the

*Applied Acoustics Laboratory, School of Electrical Engineering, Seoul National University
Manuscript Received : January 14, 2000.

subjects were amateur musicians. 18 subjects were in their twenties and the other 2 subjects were in their thirties.

2.2. Material

The original was a violin tone recorded in an anechoic chamber monaurally with 48000 16-bit samples per second. Among the tones played by an expert, the tone of A4 note (440Hz) without vibrato was selected as in Fig. 1, of which the duration was 2.7 seconds. The tone has about 40 harmonics and the spectrum is generally stable throughout time.

To consider both the steady state and the time-varying factors, 45 test tones in Tab. 1 were obtained by modifying the original violin tone. The modified tones are categorized as four groups.

In the first group, the spectrum envelopes were changed. Frequency selective filters were used for tones from 1 to 14 and from 31 to 34. From 29 to 30 smooth lowpass filters were used to change the decay slope of spectrum envelope that was originally about 3dB/oct.

In the second group, the inharmonicity was considered. Inharmonic sinusoids were added to tones from 23 to 25. Some specific harmonics were deleted in tones from 27 to 28 and from 37 to 39. Specific harmonics of tones from 40 to 45 were shifted in the frequency domain by bandpass filtering and modulation.

In the third group, the time-varying characteristics of the tones were changed. Tones from 15 to 22 were filtered only in attack or only in steady state. Tones from 35 to 36 were time-varying filtered: the bandwidth of the filter for the tone 35 was gradually increased and that for the tone 36 was decreased.

In the last group, the tone 26 was time-reversed, which approximately means that the attack and the release were interchanged.

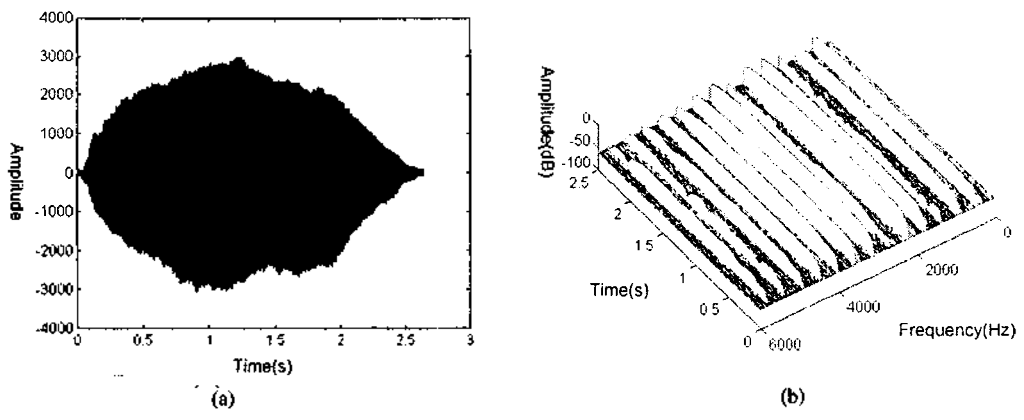
Since one modification can change the several characteristics of the tone, good care was taken of the procedure. Besides, finite impulse response (FIR) filters were used to preserve the phase characteristics of the tones. Since the largest filter tab was 2048, the modified tones have duration of less than 2.75 seconds. The energies of the tones were set to 70dB and pitches of the test tones were equalized except the tones from 40 to 45 whose modifications inevitably resulted in some pitch change. In the experiment, listeners were asked to disregard the pitch difference, if any, between the original and test tone, in consideration of tones from 40 to 45.

2.3. Apparatus

The original tone recorded in DAT was transferred digitally via optical link to a personal computer, and then modified into 45 test tones using Matlab. Pairs of the original tone and each modified tone were digitally recorded in DAT. Subjects listened to the test tones through high quality headphones.

2.4. Procedure

Authors collected a number of adjectives describing the timbre change from the previous researches [2,6,7]. Using the original and 45 modified violin tone pairs, the pretest was performed for four subjects with 30 adjective pairs. The subjects responded that they had much difficulty in filling in the questionnaire due to the large size of the adjective pairs. According to the result of the pretest, the



(a) (b)

Figure 1. Waveform and STFT spectrum of the original violin tone; (a) waveform, (b) STFT.

Table 1. The test tones made from the original violin tone (fc : cutoff frequency).

No.	Modification	No.	Modification
1	Lowpass filtered (f_c : 13400Hz)	27	4*nth harmonics deleted(n=1,2,...)
2	Lowpass filtered (f_c : 9000Hz)	28	9*nth harmonics deleted(n=1,2,...)
3	Lowpass filtered (f_c : 4600Hz)	29	Decay slope of spectrum envelope changed
4	Lowpass filtered (f_c : 2400Hz)	30	(6dB/oct)
5	Highpass filtered (f_c : 600Hz)	31	Decay slope of spectrum envelope changed
6	Highpass filtered (f_c : 1000Hz)	32	(2dB/oct)
7	Highpass filtered (f_c : 1900Hz)	33	Smoothly filtered (peak at 440Hz)
8	Highpass filtered (f_c : 4600Hz)	34	Smoothly filtered (peak at 3kHz)
9	Bandpass filtered (f_c : 600, 3300Hz)	35	Smoothly filtered (peak at 6kHz)
10	Bandpass filtered (f_c : 3300, 9000Hz)		Smoothly filtered (peaks at 1.8kHz & 3kHz)
11	Bandpass filtered (f_c : 1100, 11200Hz)	36	Time-varying Lowpass filtered (initial f_c : 4.6kHz,
12	Bandstop filtered (f_c : 1500, 2900Hz)		final f_c : 24kHz)
13	Bandstop filtered (f_c : 4200, 6800Hz)	37	Time-varying Lowpass filtered (initial f_c : 24kHz,
14	Bandstop filtered (f_c : 1100, 6800Hz)	38	final f_c : 4.6kHz)
15	Lowpass filtered in attack (f_c : 4600Hz)	39	Even harmonics deleted
16	Lowpass filtered in steady state (f_c : 4600Hz)		Odd harmonics except the fundamental deleted
17	Highpass filtered in attack (f_c : 1900Hz)	40	Odd harmonics except the fundamental and the
18	Highpass filtered in steady state (f_c : 1900Hz)		third deleted
19	Bandpass filtered in attack (f_c : 1900, 9000Hz)	41	All harmonics shifted in the frequency domain
20	Bandpass filtered in steady state (f_c : 1900, 9000Hz)		(100Hz)
21	Bandstop filtered in attack (f_c : 1900, 4600Hz)	42	All harmonics shifted in the frequency domain
22	Bandstop filtered in steady state (f_c : 1900, 4600Hz)	43	(-100Hz)
23	Inharmonic sinusoidal tones added (from f_1 to f_5)	44	The fundamental shifted (100Hz)
24	Inharmonic sinusoidal tones added (from f_{10} to f_{20})	.	The fundamental shifted (-100Hz)
25	Inharmonic sinusoidal tones added (from f_{20} to f_{30})	45	Odd harmonics shifted 100Hz, even harmonics
26	Reversed in the time domain		shifted -100Hz
			Even harmonics shifted 100Hz, odd harmonics
			shifted -100Hz

14 adjective pairs not frequently marked were discarded and the 16 adjective pairs shown in Tab. 2 were finally selected.

In the main test the subjects listened to the original tone and the modified tone successively, and then expressed timbre change in terms of the 16 adjectives by a 7-point-scale method. The subjects listened to the tones more than three times in a random order. The interval between the original tone and the modified tone was 4 seconds. The whole experiment lasted for about 40 minutes.

Table 2. Pairs of adjectives to describe the timbre change.

No.	adjective	No.	adjective
1	tenor-bass	9	light-heavy
2	clear-unclear	10	sharp-dull
3	free-pressed	11	smooth-rough
4	easy-uneasy	12	deep-shallow
5	bright-dark	13	rich-thin
6	warm-cold	14	strong-weak
7	full-hollow	15	reverberant-dead
8	broad-narrow	16	solid-split

2.5. Results

Factor analysis[8] was used to reduce the dimensions of psychological timbre factors of 45 test tones. Since we had 3-dimensional data of adjectives, 45 test tones and 20 subjects, subjects' answers were averaged to make the data 2-dimensional. The method of factor extraction was the iterated principal axis factoring with initial priors one, and four factors whose eigenvalues were 4.6089, 3.6452, 2.0407, 0.6841 respectively were retained by minimum eigenvalue method. According to Tab. 3 which shows the rotated factors by varimax method, the first factor is characterized by the adjectives, 'free', 'broad', 'deep', 'rich', 'strong', and 'reverberant', the second by 'tenor','clear', 'bright', 'light', and 'sharp', the third by 'easy', 'smooth', and 'solid' and the fourth by 'warm' and 'full'. The final rotated factors explain 39.3%, 30.9%, 17.7%, 12.1% of the common variance respectively, and the common factors explains 68.6% of the total variance. Correlation coefficients in the same factor ranged from 0.4359($p<0.0028$) to 0.8346($p<0.0001$). For detailed values of correlation coefficients, refer to [9].

Table 3. Factor loadings of 4 rotated factors and communalities.

Adjectives	F1'	F2'	F3'	F4'	communality
Tenor-bass	-0.09175	0.71873	-0.28635	-0.01984	0.607377
clear-unclear	0.05190	0.83968	0.08189	0.05073	0.717031
free-pressed	0.83328	0.39522	-0.09159	-0.03962	0.860510
easy-uneasy	-0.07242	-0.11860	0.80200	0.13105	0.679694
bright-dark	-0.02310	0.72351	-0.03567	0.40978	0.693195
warm-cold	0.08120	0.03331	0.30721	0.60625	0.469618
full-hollow	0.15595	0.00578	0.18279	0.71522	0.569307
broad-narrow	0.87055	-0.12130	-0.20725	0.17488	0.846100
light-heavy	-0.26601	0.71636	-0.00892	-0.14613	0.605370
sharp-dull	-0.02422	0.86974	-0.36421	-0.10747	0.901239
smooth-rough	-0.13675	-0.06843	0.53888	0.25570	0.379153
deep-shallow	0.81765	-0.28809	0.10546	0.14851	0.784731
rich-thin	0.84536	-0.25581	-0.05718	0.22796	0.835302
strong-weak	0.86162	-0.11261	-0.14558	0.12043	0.790766
reverberant-dead	0.76158	0.06757	-0.08021	-0.18279	0.624413
solid-split	-0.12524	-0.13625	0.75322	0.11594	0.615034
Variance	4.311317	3.389772	1.944794	1.332955	total communalities : 10.9788

III. Discussion

The first factor seems to be related to 'richness' of a tone, which is a good contrast to Bismarck's research[2] in which 'sharpness' played the most important role in steady state tones. Due to the effects of the time-varying modifications, the 'richness' or the 'colorfulness' explains the largest variance of the timbre space.

The second factor is related to the already known timbre of 'sharpness' that depends on the energy centroid of a tone. It shows that sharpness plays an important role in a time-varying tone as well as in a steady state tone. In particular, tones from 35 to 36 were misleadingly perceived as pitch glide and, what is more, listeners perceived tones from 15 to 22 as successive two tones of different pitch.

The third and fourth factors can be named as 'smoothness' and 'fullness' As in Bismarck's research[2], 'fullness' acted as distinct timbre space from others.

IV. CONCLUSION

4 factors were extracted with the timbre of 45 modified test tones through factor analysis. 'Richness', 'sharpness','smoothness' and 'fullness' can explain each factor respectively. The first factor, 'richness', seems to be dependent upon the time-varying characteristic of a tone. The second factor, 'sharpness', plays an important role in a time-varying tone as well as in a steady state tone and shows some evidence of the fact that time-varying sharpness exists.

Since both physical and psychological factors of timbre have multi-dimensional characteristics, some restriction such as dimension reduction of either physical or psychological factors, we think, is strongly required in timbre research. Moreover, timbre which one feels in music has yet to be studied since it creates another psychological timbre space such as 'sad' or 'exciting' that is hardly found when one hears a single tone.

References

1. J. M. Grey, An exploration of musical timbre, Ph. D. dissertation, Stanford Univ., 1975.
2. G. von Bismarck, "Timbre of steady sounds: A factorial investigation of its verbal attributes", Acustica, **30**, pp. 146-159, 1974.
3. G. von Bismarck, "sharpness as an attribute of the timbre of steady sounds", Acustica, **30**, pp. 159-172, 1974.
4. J. M. Grey, "Multidimensional perceptual scaling of musical timbres," J. Acoust. Soc. Am., **61**, pp. 1270-1277, 1977.
5. P. Iverson, C. L. Krumhansl, "Isolating the dynamic attributes of musical timbre," J. Acoust. Soc. Am., **94**, pp. 2595-2603, 1993.
6. G. Bloothooft, R. Plomp, "The timbre of sung vowels," J. Acoust. Soc. Am., **84**, pp. 847-860, 1988.
7. R. Ethington, B. Punch, "Sea Wave : A system for musical timbre description," Computer Music Journal,

18, pp. 30-39, 1994.
8. S. A. Mulaik, The foundations of factor analysis, McGraw-Hill, New York, 1972.
9. Hee-Suk Pang, Chul-Yong Ahn, Koeng-Mo Sung, "Subjective timbre space of violin tones modified by various methods," Proc. of WESTPRAC VI, 1, pp. 197-202, 1997.

▲ Chul-Yong Ahn

Chul-Yong Ahn received the BS and MS degrees in Electronics Engineering from Seoul National University in 1993 and 1995, respectively. Now he is working toward the Ph. D. degree at Seoul National University. His current research interests are in the area of digital signal processing, psychoacoustics, multi channel sound field reproduction and binaural recording.

▲ Hee-Suk Pang

Hee-Suk Pang received the BS and MS degrees in Electronics Engineering from Seoul National University in 1994 and 1996, respectively. Now he is working toward the Ph. D. degree at Seoul National University. His current research interests are in the area of digital signal processing, psychoacoustics and musical acoustics.

▲ Koeng-Mo Sung

The Journal of the Acoustical Society of Korea, Vol. 17. No. 3E, 1998.