PSYC 562 Measurement of Psychological Processes

Final Project

Dissimilarity of synthesized sounds as a function of timbral descriptors

10 December 2008

Song Hui Chon

Abstract

A multidimensional scaling analysis was applied to a set of dissimilarity data, obtained from 20 subjects on 20 synthetic stimuli, in two different criteria (maximum likelihood and individual differences) to obtain 3-dimensional solutions. An iterative correlation analysis was applied to the MDS solutions to find the rotational direction of the two-dimensional planes that maximize the correlations between an MDS dimension and a design parameter. As the result shows, two of the three 2-dimensional planes in the 3-dimensional MDS spaces can be rotated to certain degrees so that each of the dimensions will have a maximum correlation with a design parameter. This seems to suggest a one-to-one mapping between an MDS dimension and a design parameter, and furthermore an orthogonal relationship among the design parameters in timbre space. If this hypothetical one-to-one mapping is true, this will validate the three-dimensional timbre space model with three axes of attack time, spectral centroid and spectral fluctuation, all of which are given as design parameters in this paper.

1. INTRODUCTION

A sound can supposedly be described by a number of standard descriptors, such as pitch, loudness and timbre. Among those descriptors *timbre* is the least understood so far, simply because of its multi-faceted nature. A standard definition of *timbre* is "[...] that attribute of sensation in terms of which a listener can judge that two sounds having the same loudness and pitch are dissimilar" [1]. This sentence means that *timbre* can be used for any characteristics of sounds that are different and yet have the same loudness and pitch.

Even though this multi-dimensionality of *timbre* makes it hard to study, there has been a growth of interest in understanding and defining the structure of timbre. Multidimensional scaling has been a popular technique to describe the multi-dimensionality of timbre and 3-dimensional solutions have been particularly popular ([2][3][4][5]) due to its potential interpretability and predictability. So far, there is no standard 3-dimensional timbre space model that everyone agrees on, but most agree that the two most important dimensions are highly correlated with attack time and spectral centroid (or spectral center of gravity). The third dimension may have less number of consensuses than the other two dimensions, but this paper assumes (and later verifies) that this dimension is highly correlated with the spectral fluctuation, following the model by McAdams et al. [6]

2. STIMULI

Table 1: Parameter values for each stimulus

Stimulus	Attack	SCG	EHA
1	29.61794	3	0
2	25.8501	3.078947	7.157895
3	87.96233	3.157895	4.631579
4	51.04178	3.236842	6.736842
5	58.4815	3.315789	1.263158
6	33.93498	3.394737	5.894737
7	100.7835	3.473684	6.315789
8	115.4735	3.552632	1.684211
9	199	3.631579	2.105263
10	76.7722	3.710526	4.210526
11	173.6842	3.789474	8
12	19.69141	3.868421	5.473684
13	67.00562	3.947368	2.526316
14	17.18636	4.026316	2.947368
15	22.56158	4.105263	7.578947
16	151.589	4.184211	3.368421
17	44.5485	4.263158	0.8421053
18	132.3046	4.342105	5.052632
19	38.88126	4.421053	0.4210526
20	15	4.5	3.789474

Twenty sounds were generated following the model by Caclin, et al. [7] Three parameters – Attack, SCG (Spectral Center of Gravity) and EHA (Even Harmonic Attenuation) – were manipulated to generate synthetic stimuli. Table 1 shows the values of the three parameters for each stimulus.

Attack corresponds to the duration of the attack time for a stimulus, in millisecond. The smaller the attack value is the sharper the attack is, hence the sharper the perceived timbre will be. SCG is an amplitude-weighted mean frequency of the energy spectrum (in terms of the harmonic index) and EHA is the degree of a selective attenuation of even harmonics relative to odd harmonics, in the range of 0 to 8 dB.

Table 2: Sorting orders of stimuli according to each parameter

Stimulus	Attack	SCG	EHA
1	20	1	1
2	14	2	19
3	12	3	17
4	15	4	5
5	2	5	8
6	1	6	9
7	6	7	13
8	19	8	14
9	17	9	16
10	4	10	20
11	5	11	10
12	13	12	3
13	10	13	18
14	3	14	12
15	7	15	6
16	8	16	7
17	18	17	4
18	16	18	2
19	11	19	15
20	9	20	11

The orderings of the 20 stimuli in an ascending order according to each of the three parameters are presented in table 2. These orderings will be referred to as “indices” (e.g., “EHA indices”) and used in the section 5. Results and Discussions.

3. SUBJECTS

Twenty subjects with normal hearing participated in this experiment. Specifics on the subjects’ biographical background are unknown at this time, although they must have been collected during the experiment.

4. DESCRIPTION OF DATA

The subjects first listened to all 20 stimuli multiple times to get familiar with the overall range of dissimilarity among the stimuli. They were then asked to judge the dissimilarity of a given pair of stimuli using a controller, whose range was mapped from 0 to 1. 0 meant the two sounds were perceptually identical and 1 meant the two sounds were as different as possible. Each subject was presented with all 190 pairs (the number of possible unique pairs with 20 stimuli). The author of this paper does not know the specifics of the procedure and setup for the experiment.

5. RESULTS & DISCUSSION

The MULTISCL program was used for the multidimensional scaling analysis. Two types of analyses were carried out, one with ML (maximum likelihood) criterion and the other with ID (individual differences) criterion. In both cases, the best solutions turn out to be the three dimensional ones.

An iterative correlation analysis was applied on the 3-dimensional MDS solutions for both ML and ID cases. The purpose of this part was to find the direction that maximizes the correlation between a dimension in an MDS solution and a design parameter (i.e., Attack, SCG or EHA).

5. 1. *Iterative Correlation Analysis*

With a 3-dimensional MDS solution given, each of three planes (formed by axes 1 & 2, by axes 1 & 3, and by axes 2 & 3) was rotated from 1 degree to 359 degrees by 1-degree increment. The correlation coefficients were calculated between each pair of the three MDS dimensions and the three parameters per rotation. Then the maximum correlation coefficients were found for each pair of the MDS dimensions and the parameters. The angles of rotation corresponding to the maximum correlation coefficients then reflect the directions of the three design parameters embedded in the MDS space.

The analysis results are presented in the tables 3 and 4. For the ML MDS solution in table 3, the maximum correlation of 0.9085 was found between the dim3 and Attack after a 334-degree (clockwise) rotation of the dim2–dim3 plane. The rotation of dim1–dim2 plane by 224 degrees yielded the maximum correlation of 0.9161 between the dim2 and the SCG parameter as well as the maximum correlation of 0.8605 between the dim1 and the EHA parameter. Similarly for the ID MDS solution presented in table 4, the Attack parameter showed the highest correlation with dim3 after a rotation of the dim2–dim3 plane. However, the other two parameters showed the highest correlations with different dimensions after a rotation of the same (dim1–dim2) plane, when compared with the ML case in table 3. This seems to suggest that the Attack parameter is a function that can be mostly described by dim3 of the MDS dimension in either of ML or ID criterion, while the other two parameters (SCG and EHA) may be functions of a combination of both dimensions 1 and 2 and that may be why different criterion produces different mappings as we observed. It is also possible that the dimensions 2 and 3 are swapped for ML and ID solutions and each of those dimensions is indeed a function of one parameter.

Table 3: Correlation analysis result for the ML MDS solution

Parameter	Maximum Correlation	Corresponding axes after rotation of plane
Attack	0.9085	dim3 after rotating dim2–dim3 plane by 334 degrees
SCG	0.8601	dim1 after rotating dim1–dim2 plane by 224 degrees
EHA	0.9161	dim2 after rotating dim1–dim2 plane by 224 degrees

Table 4: Correlation analysis result for the ID MDS solution

Parameter	Maximum Correlation	Corresponding axes after rotation of plane
Attack	0.8962	dim3 after rotating dim2–dim3 plane by 341 degrees
SCG	0.9339	dim2 after rotating dim1–dim2 plane by 341 degrees
EHA	0.9397	dim1 after rotating dim1–dim2 plane by 18 degrees

Even though this method does work (judging from the large values of maximum correlation coefficients), it has a couple of serious limitations. First, this assumes that there is a one-to-one mapping between each of the MDS dimensions and the three parameters. This may be true but it may also be too idealistic.

Second, this approach does not guarantee a unique solution. Regression analysis will provide a unique solution for placing the three parameters in the MDS space; however, due to the limitations of time and the lack of experience with this analysis method, the author decided to use the method described above.

5. 2. *ML MDS Result*

Figure 1 shows the 3-dimensional solution using the ML criterion. Detailed plots of the solution on three 2-dimensional planes are presented in figures 2 – 4.

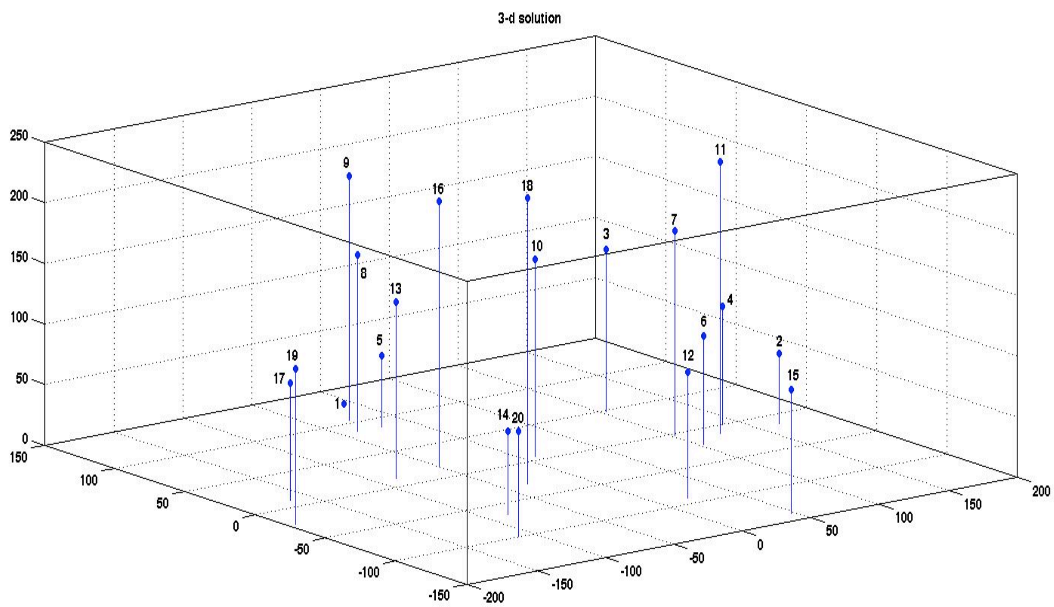


Figure 1: 3-D ML MDS solution

Figure 2 shows the dim1—dim2 plane twice; the one on the left shows the grouping by EHA indices and the one on the right shows the grouping by SCG indices. An EHA index is the position of the particular stimulus's EHA value, where all 20 EHA values are sorted in an ascending order. Similarly, SCG and Attack indices are used for the respective positions of SCG and Attack values. These indices are presented in table 2.

The red arrows in the two figures reflect the dimensions 1 and 2 after a rotation of 224 degrees, which respectively produces the maximum correlations with SCG and EHA parameters, as explained in the last section.

From the figure on the left of figure 2, we can see that the EHA index values tend to increase along the direction of dim2 (pointed by a red arrow). The figure on the right of figure 2 illustrates that the SCG values seem to increase in the same direction of dim1. These visual patterns confirm the correlation analysis result on table 3.

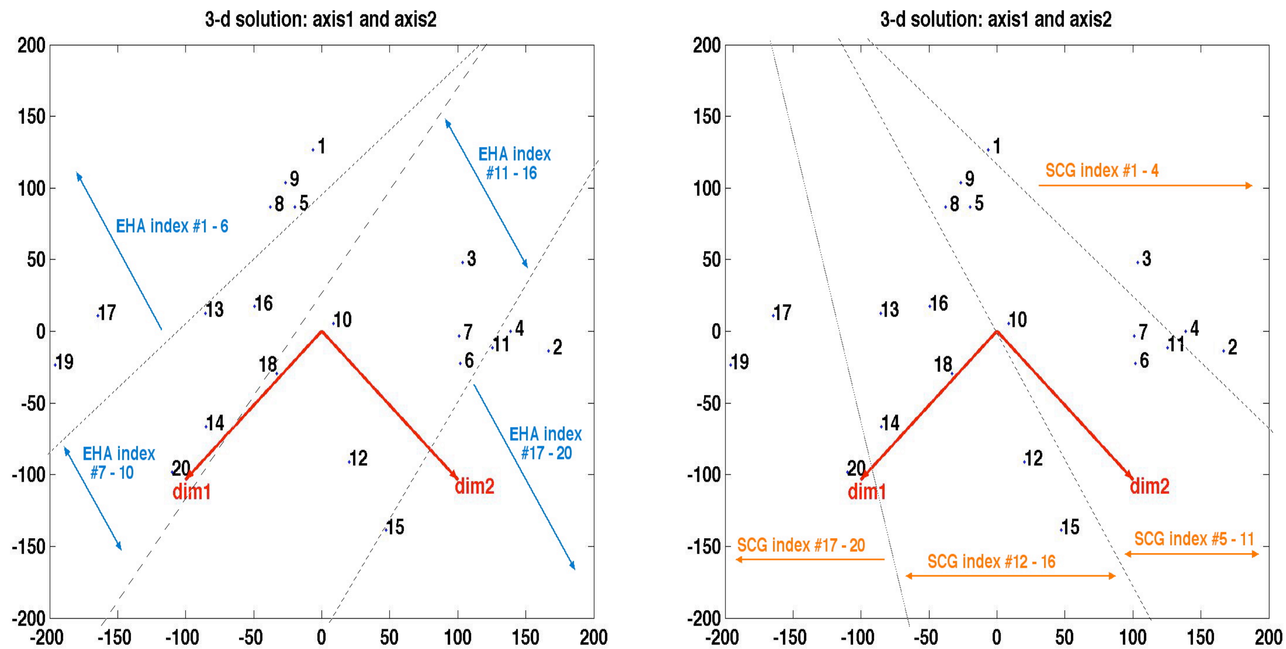


Figure 2: Dim-1 & 2 of the 3-D ML MDS solution

Figure 3 shows the dimensions 1 & 3 of the ML MDS solution. We can see that the Attack index tends to increase along the direction of dim3. This appears to support the correlation analysis result in table 3, although the visible correlations are a lot weaker than in figure 2 (for example, see the "Attack index 1—11" group). Even though a visible division can be observed in the stimuli configuration with EHA values in figure 3, the correlation between the EHA parameter and the first dimension seems negligible. The plane was not rotated at all, since the correlation analysis did not find a maximum correlation between a dimension and a parameter on this plane.

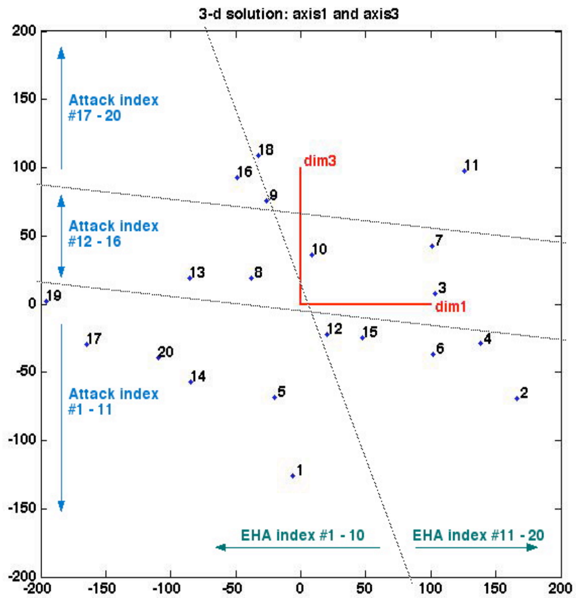


Figure 3: Dim-1 & 3 of ML MDS solution

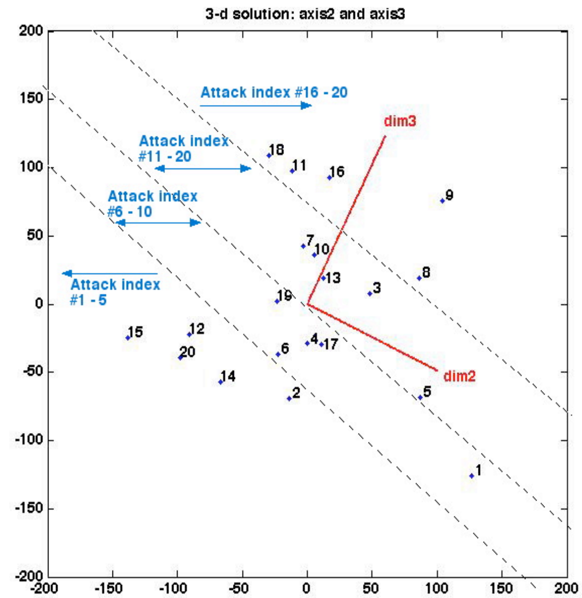


Figure 4: Dim-2 & 3 of ML MDS solution

The plane formed by dimensions 2 & 3 is illustrated in figure 4. The dimensions are shown with a 334-degree rotation. Here we can observe that the Attack index increases with the same direction of dim3. This is in agreement with the correlation analysis result in table 3. No other visual patterns of correlations were detected on this plane.

5. 3. *ID MDS Result*

Figure 5 shows the configuration of 20 stimuli using the ID criterion. Comparing with the ML MDS solutions in figure 1, there seems to be similar local groupings but at the same time it is obvious that the either solution cannot be obtained from a simple manipulation of the other, such as by rotating or mirror imaging.

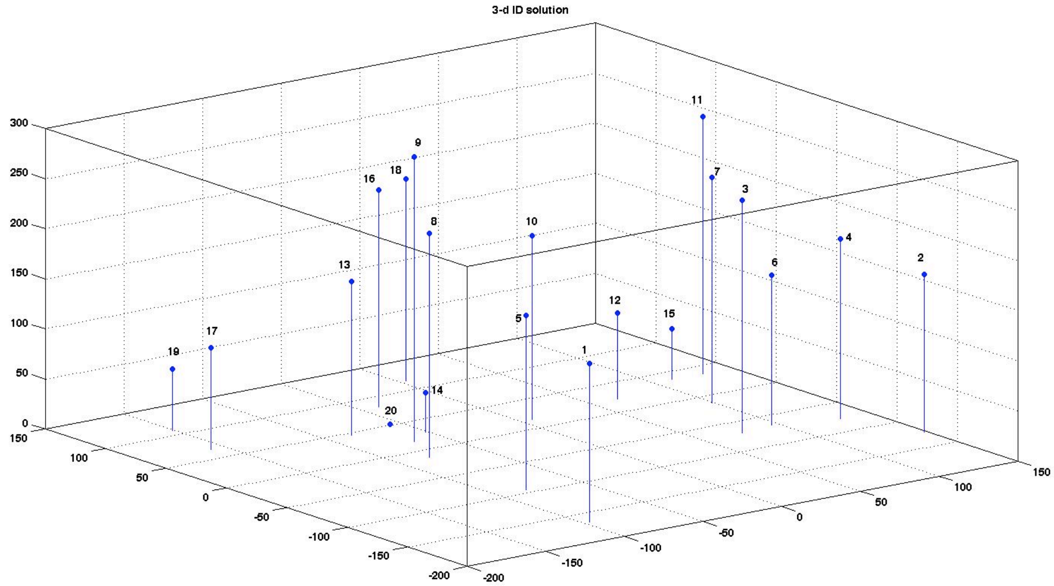


Figure 5: 3-D ID MDS solution

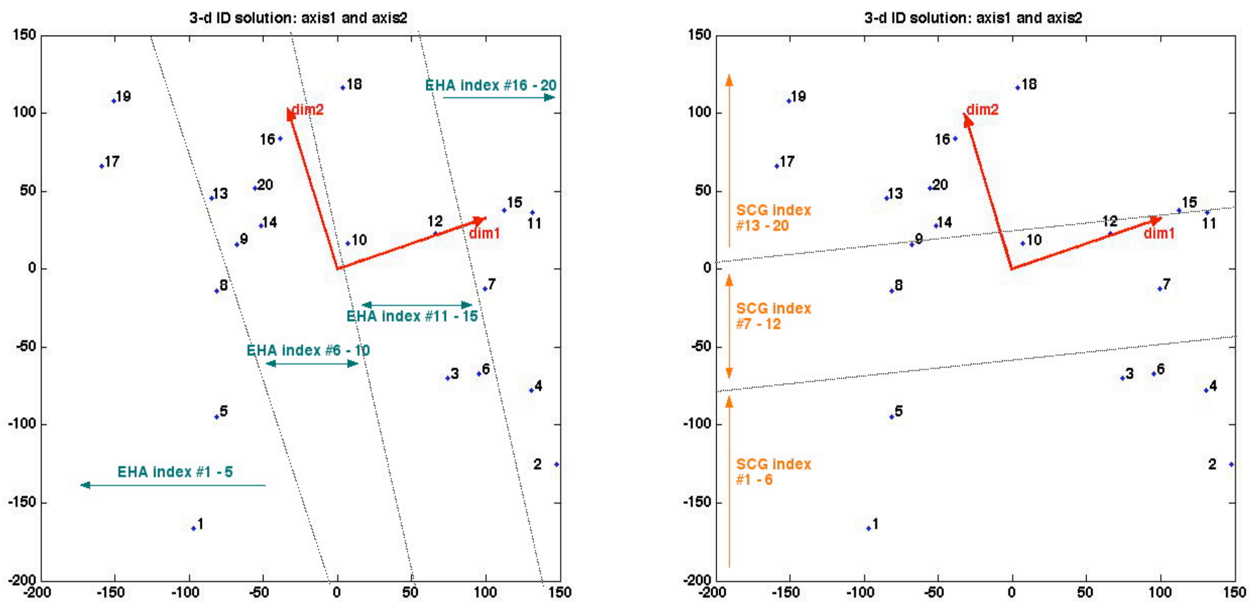


Figure 6: Axes 1 & 2 of ID MDS solutions

The dim1—dim2 plane is illustrated twice in figure 6. The two red arrows show the direction of both dimensions after a rotation of 18 degrees (as explained in the table 4). The figure on the left shows that the EHA index seems to be increasing along the direction of dim1. The figure on the right shows that the SCG index increases along the direction of dim2. These visible patterns again confirm the correlation analysis result in table 4.

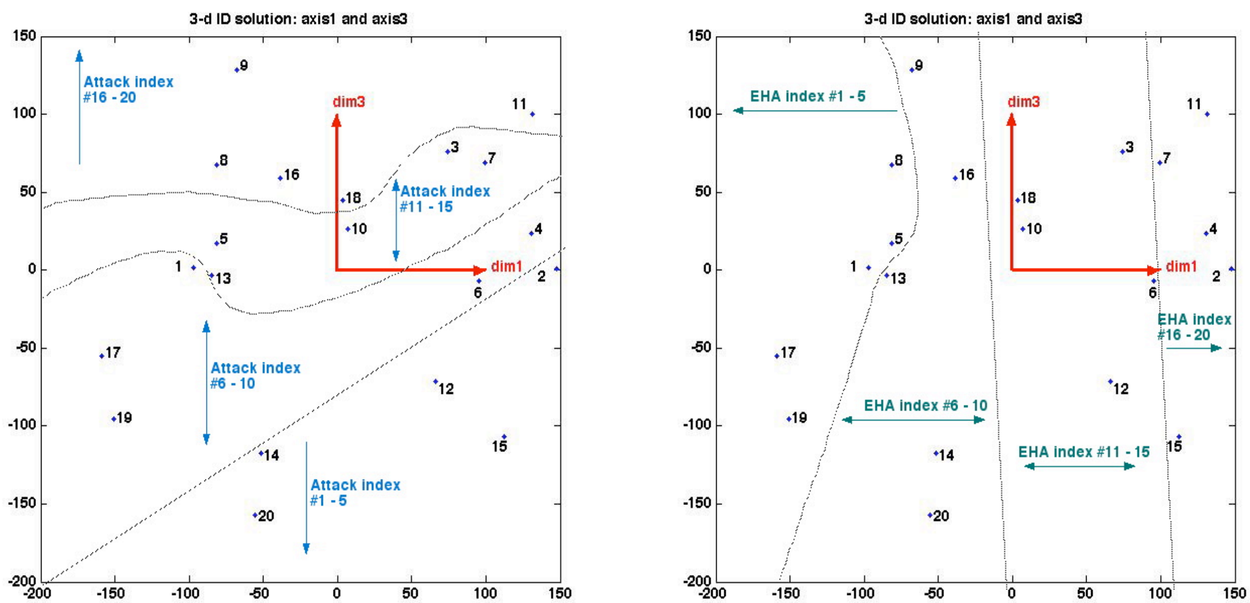


Figure 7: Axes 1 & 3 of ID MDS solution

Figure 7 shows the dim1—dim3 plane of the ID MDS solution. This plane was not rotated, with the same reason as in the ML MDS case, presented in figure 3. On the left, we can see that the Attack index seems to increase with the dim3 direction, although the patterns seem to be a bit weaker than other cases. The figure on the right shows that the dim1 seems to be correlated with the EHA indices.

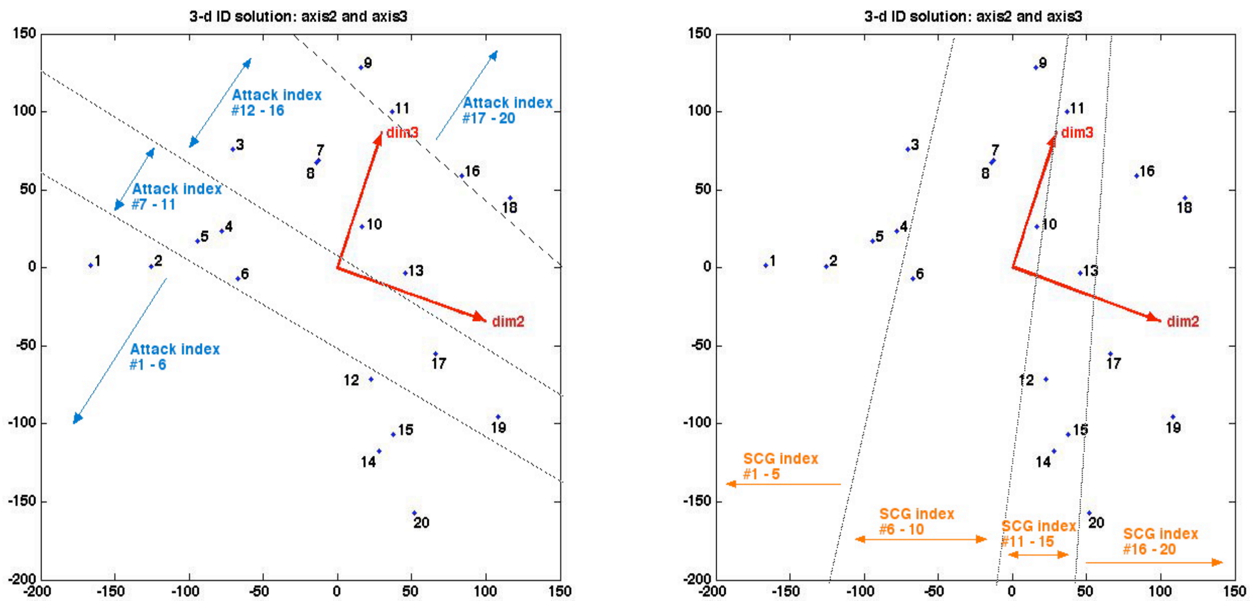


Figure 8: Axes 2 & 3 of ID MDS solution

Figure 8 presents the dim2—dim3 plane of the ID MDS solution. The red arrows show the directions of the two dimensions after a 341-degree rotation. Here we can observe a cleaner pattern than in figure 7; the figure on the left shows that the Attack index tends to increase along the direction of dim3, and the figure on the right shows that the SCG index seems to grow following dim2. These patterns once again visually confirm the correlation analysis result.

The individual weights associated with the ID MDS solution were illustrated in figures 9 and 10. Figure 9 seems to suggest that there may be different clusters in individual weights. Interestingly, there is a void around (1, 1, 1) in the constellation. Since the points around (1, 1, 1) will correspond to the equal contribution of the three dimensions in describing the given data, this void seems to mean that no subject used an equal combination of the three dimensions in the dissimilarity judgment of the twenty stimuli. A further analysis with more subjects is required for the verification of this possible clustering pattern in individual weights.

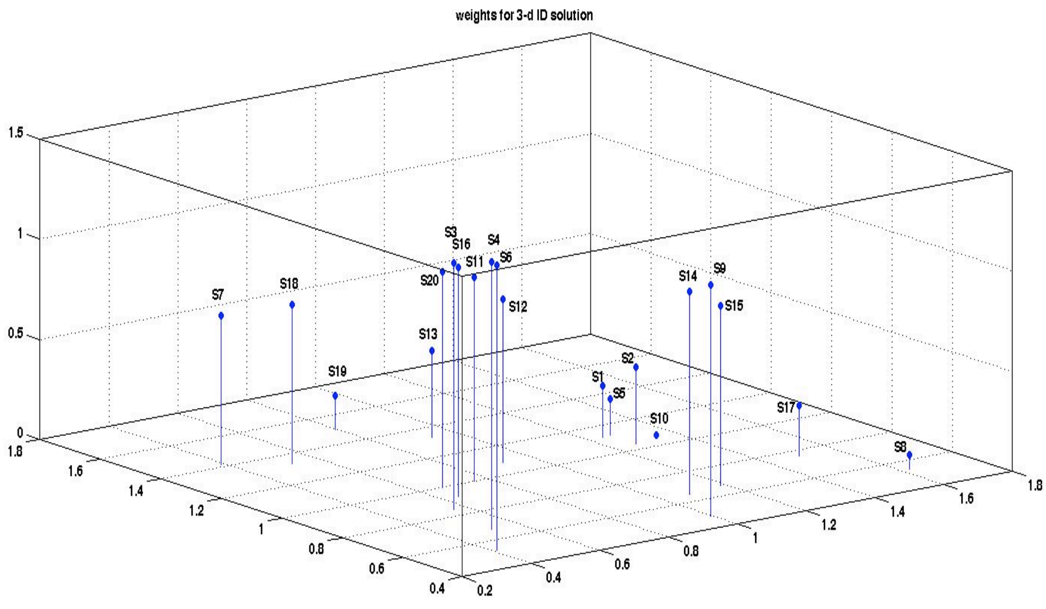
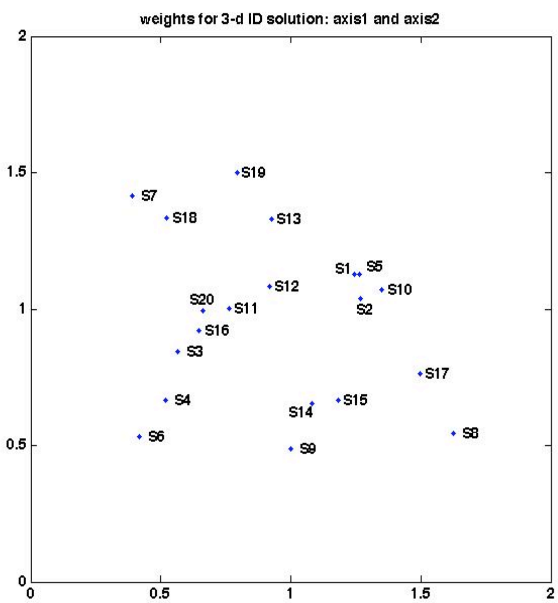
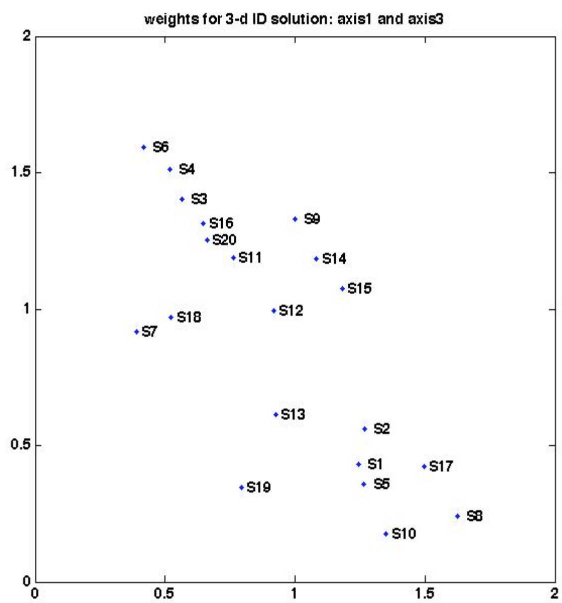


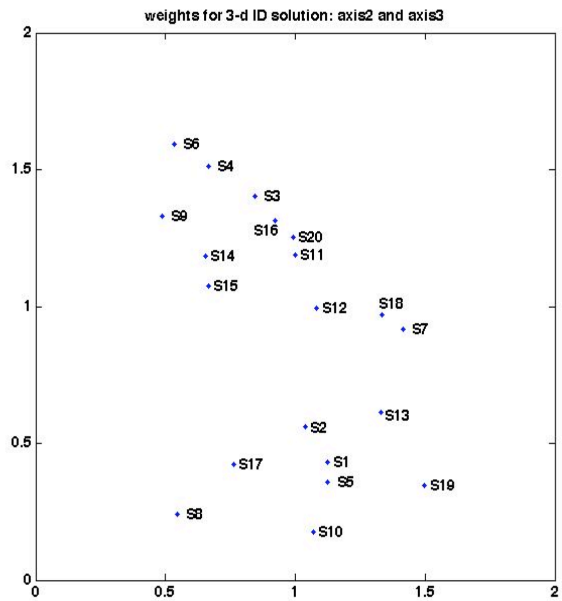
Figure 9: Individual weights for 3-D ID solutions



(a) Axes 1 & 2



(b) Axes 1 & 3



(c) Axes 2 & 3

Figure 10: Individual weights on three planes in 3-D ID MDS solution

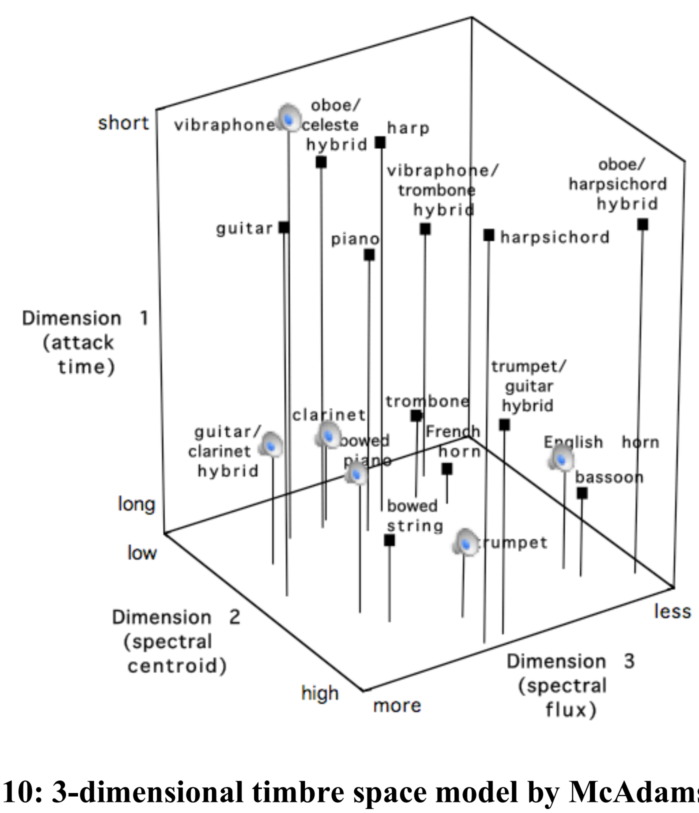


Figure 10: 3-dimensional timbre space model by McAdams et al. [6]

6. CONCLUSION

This paper investigated the relationship of dissimilarity ratings with three acoustic correlates – Attack, SCG and EHA. The three parameters are important in understanding timbre space, since they are considered to form orthogonal dimensions in a 3-dimensional MDS timbre space, as shown in figure 10 (reproduced from [6]).

MDS solutions were obtained using two different criteria of ML and ID. In both cases, the three dimensional solutions were found to be the optimal.

An iterative correlation analysis was applied to the MDS solutions to find the directions of rotations of three planes that maximize the correlations between each MDS dimension and each parameter. The analysis revealed that there is a set of rotational angles that maximizes one-to-one correspondence between the MDS dimensions and the parameters. Since the MDS dimensions are orthogonal by definition, this one-to-one correspondence implies that the three parameters are orthogonal to one another, therefore successfully forming a three-dimensional space. This gives credibility to the three-dimensional timbre space model using the same three acoustic correlates in timbre research. This is remarkable since the subjects were not asked to judge dissimilarity of a pair of stimuli in a specific category (related to the design parameter). But with the correlation analysis on the three dimensional solution, we can say that the three parameters are sufficient to successfully describe the timbral dissimilarity of sound stimuli.

The iterative correlation analysis has a couple of serious limitations, due to the over-simplified assumption of one-to-one mapping between a MDS dimension and an acoustic correlate, as well as no guarantee of a unique solution. A further study using regression analysis will be necessary to find a unique solution for this problem.

For the ID MDS solution, the 3-dimensional weight plot seems to suggest that there may be clusters in the patterns of timbre perception. It is worth a notice that none of the subjects put the same weight on all three dimensions. A cluster may reveal one dimension more preferred to the other two in timbre perception, which may depend on the subjects' biographical data. A more thorough analysis is required with more subjects to verify this conjecture.

REFERENCE

[1] American National Standard Institute (1973). *American national psychoacoustical terminology.* S3.20. New York: American Standards Association.

[2] Grey, J. M., & Gordon, J. W. (1978). Perceptual effects of spectral modifications on musical timbres. *Journal of the Acoustical Society of America*, 63, 1493—1500.

[3] Ehresman & Wessel 1978

[4] Kendall, R. A., & Carterette, E. C. (1991). Perceptual scaling of simultaneous wind instrument timbres. *Music Perception*, 8, 369—404.

[5] McAdams, S., & Cunibile, J-C. (1992). Perception of timbral analogies. *Philosophical Transactions of the Royal Society, London, Series B,* 336, 383—389.

[6] McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychol Res*, 58, 177—192.

[7] Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America*, 118 (1), 471—482