Axes Rectifying of Impression Space in Music Impression-based Retrieval and its Evaluation

TSUYOSHI TAKAYAMA[1], YUUTA HASHIBA[2], SHIGERU KIKUCHI[1],
TETSUO IKEDA[3] and YOSHITOSHI MURATA[1]
[1] Graduate School of Software and Information Science
Iwate Prefectural University
152-52, Sugo, Takizawa, Iwate020-0193
JAPAN

http://www.iwate-pu.ac.jp/
[2] Hanamaki Head Office and Factory
Shinko Seisakusho Co., Ltd.
4-3, Jonai, Hanamaki, Iwate025-8502
JAPAN
http://www.shinko-exc.co.jp/
[3] School of Administration & Informatics
University of Shizuoka
52-1, Yada, Suruga-ku, Shizuoka, Shizuoka422-8526
JAPAN
http://www.u-shizuoka-ken.ac.jp/

Abstract: - Recently, lots of researchers are attracted to retrieving multi-media databases by impression[1][2]. As an example, Ikezoe *et al.* propose to retrieve multi-media databases by using eight pairs of opposite impression words[3]. In case of realizing this approach by applying Salton's vector space[4], we are required to rectify each weight of eight axes. In this paper, we propose three rectifying methods and evaluate which method is the most appropriate. Concerning the best one, we also develop a pilot system and execute a relative comparison from the conventional method.

Key-Words: - Multi-media database, music database, ambiguous retrieval, impression-based retrieval.

1 Problem Description

The Ikezoe's music impression space[3] is multi-dimensional, and it has eight pairs of opposite impression words. Each axis has seven levels from minus three to plus three(Fig. 1). It is based on the SD(Semantic Differential) method[5], which represents an object by the combination of each value on multiple axes of opposite impression words. A music is placed in the space according to each value. A retrieval condition is represented as a retrieval vector, and the corresponding retrieval result is produced by neighborhood retrieval[6].

This approach has the following problem: even if a user has an image of a certain music and provides the retrieval vector for it, the music is not always emerged in upper rank of the retrieval result.

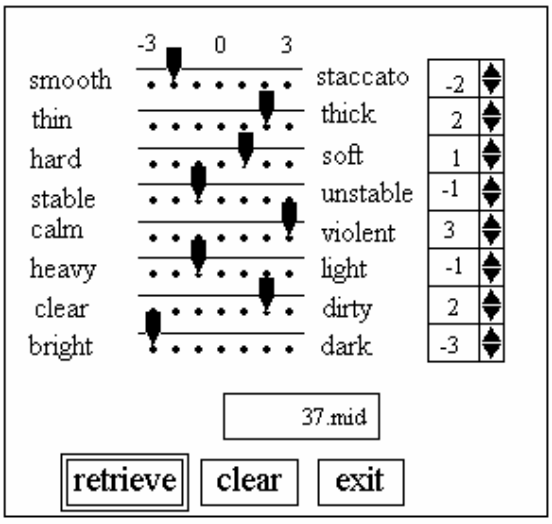

Fig. 1 Ikezoe's retrieval interface.

2 Solution

In this paper, as axes rectifying by providing appropriate weight to each axis in calculating distance for neighborhood retrieval, we propose the following three methods. Hereafter, we abbreviate 'smooth versus staccato' in Fig.1 to axis 1, and 'thin versus thick' to axis 2, etc.

Method 1: axes rectifying based on dispersion of impression value by subject for the same music(Fig. 2). This method sets a light weight to an axis whose dispersion of impression value by subject is not small. On the other hand, it sets a heavy weight to an axis whose dispersion is not big. We evaluate the dispersion by variance.

Method 2: axes rectifying based on dispersion of impression value by music(Fig. 3). This method sets a heavy weight to an axis whose dispersion of impression value by music is not small. On the other hand, it sets a light weight to an axis whose dispersion is not big. We evaluate the dispersion by variance.

Method 3: axes rectifying based on personal gap from median of impression values on an axis collected from some subjects(Fig. 4). Hereafter, we call this gap 'offset'. In this method, we set inherent weights for each user. His/her offset on an axis is determined by distance from the median to his/her impression value. If the offset is not small, we set a light weight to the axis. On the other hand, if the offset is not big, we set a heavy weight to the axis. However, in case of applying this method to his/her actual retrieval operation, it is not realistic to enforce him/her to provide impression value to all music in a music database[7]. Therefore in this paper, we select ten music as a sample, and determine the offset.

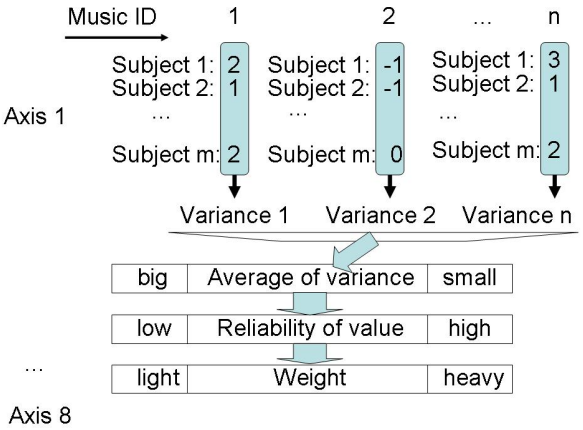

Fig. 2 Method 1.

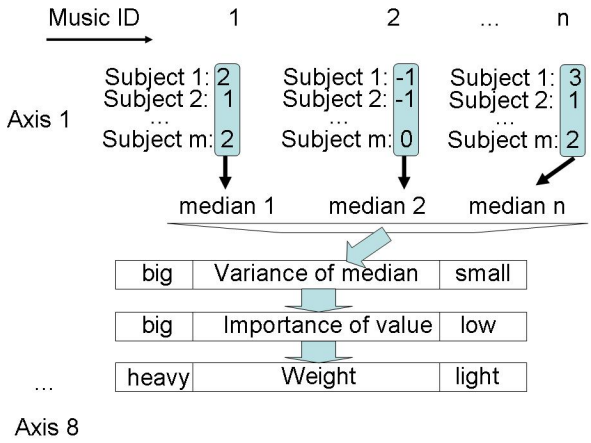

Fig. 3 Method 2.

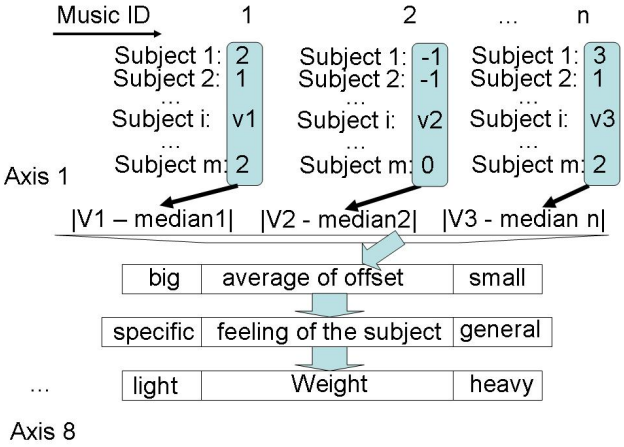

Fig. 4 Method 3.

Each default weight of eight axes is one, and the sum of eight weights is eight. Although each weight is changed by method 1, 2, or 3, we always keep the sum of eight weights to eight.

3 Evaluation

3.1 Calculating Weights and its Analysis

Using 280 Japanese popular music, we calculate each weight of eight axes in case of each method.

3.1.1 Method 1

Table 1 shows each weight of eight axes. When we apply these weights to the 280 music, Euclid distance from the retrieval vector is reduced in only 42.5 % of them. This means that distance from retrieval vector is enlarged in more half music than before axes rectifying by method 1.

Table 1 Each weight by method 1

axis 1	1.016	axis 2	0.968	axis 3	1.032	axis 4	1.008
axis 5	0.984	axis 6	1.008	axis 7	0.976	axis 8	1.016

3.1.2 Method 2

Table 2 shows each weight of eight axes.

Table 2 Each weight by method 2

axis 1	0.976	axis 2	1.440	axis 3	1.560	axis 4	0.880
axis 5	0.840	axis 6	0.928	axis 7	0.800	axis 8	0.576

When we apply these weights to the 280 music, Euclid distance from the retrieval vector is reduced in only 41.1 % of them. This means that distance from retrieval vector is also enlarged in more half music than before axes rectifying by method 2.

3.1.3 Method 3

Table 3 shows each weight of eight axes for a subject A. He/she selects ten music in four ways from the 280 music. Table 4 shows in the case of a subject B.

Table 5 shows the ratio that Euclid distance from the retrieval vector is reduced when we apply each weight to the 280 music. In the subject A, the distance is reduced 56.6% at average. In the subject B, it is 59.4% at average. From the total of two subjects point of view, the distance is reduced in 58.0% music. This means that method 3 has higher possibility for improvement in ranking of retrieval result than method 1 or 2.

Table 3 Each weight for a subject A by method 3
(1): axis, (2): ten music for weight decision

(2) / (1)	Favorite ten music for the subject A	ID= 1-10	ID= 11-20	ID= 21-30
axis 1	0.936	0.970	0.998	1.042
axis 2	0.907	0.952	0.854	0.930
axis 3	1.047	1.012	1.071	1.092
axis 4	1.018	0.982	1.018	0.981
axis 5	1.047	1.048	1.044	0.900
axis 6	1.025	1.030	1.044	0.961
axis 7	0.929	1.012	0.998	1.011
axis 8	1.091	0.994	0.972	1.082

Table 4 Each weight for a subject B by method 3
(1): axis, (2): ten music for weight decision

(2) / (1)	Favorite ten music for the subject B	ID= 1-10	ID= 11-20	ID= 21-30
axis 1	0.970	1.002	0.937	1.027
axis 2	1.004	0.972	0.876	0.919
axis 3	1.074	1.033	1.082	1.042
axis 4	0.900	0.873	0.945	0.880
axis 5	0.958	1.078	1.029	1.004
axis 6	1.062	1.037	1.112	1.006
axis 7	0.981	0.957	1.013	1.012
axis 8	1.051	1.048	1.006	1.050

Table 5 Ratio of music that Euclid distance from retrieval vector is reduced by applying each weight

	Subject A	Subject B
Favorite ten music for the subject A	65.7	
Favorite ten music for the subject B		57.1
ID=1-10	61.4	59.6
ID=11-20	53.2	59.6
ID=21-30	46.1	61.4
Average per subject	56.6	59.4
Average per 2 subjects	58.0	

3.2 Introduction to Pilot System

We have developed a pilot systems based on the method 3. We carry out the axes rectifying using the weights produced by ten music which are favorite of subject A. For relative comparison, we have also prepared the conventional system by employing the part of the system in the paper[8]. In here, 'conventional system' means that all weight of eight axes is one. Both of the two systems are implemented by using Servlet/JSP[9]. We use Oracle9i[10] as DBMS.

We examine whether the method 3 is actually more effective than the conventional. We treat the case that the music which should be emerged as one of the retrieval results is ID=161. Table 6 and 7 show the result. Although it is not emerged in the retrieval results of conventional method(Table 6), it is emerged in ones by method 3 as fourth candidate(Table 7).

Table 6 Retrieval results by conventional method

Rank	ID	Music title	Artist	$(\text{distance})^2$
1	112	Sora mo toberu l	Spits	4.00
1	116	Kimi no namae	N. Maikih	4.00
1	196	Secret base	Zone	4.00
4	108	Hatenaku tsuzuk	MISIA	5.00
4	120	Hikari to kage n	K. Udo	5.00

Table 7 Retrieval results by method 3

Rank	ID	Music title	Artist	$(\text{distance})^2$
1	196	Secret base	Zone	3.90
2	116	Kimi no namae	N. Makiha	3.93
3	112	Sora mo toberu l	Spits	4.03
4	161	Okuru Kotoba	Kaientai	4.76
5	173	Sogen no hito	A. Matsuu	4.99

4 Conclusion

In this paper, we have proposed three axes rectifying methods of music impression space for impression-based retrieval. According to our evaluation, the method 3 based on personal gap from median of impression values on an axis collected from some subjects is relatively more effective than the others. We have also developed a pilot system based on the method 3, and executed a relative comparison from the conventional method. Its result has shown that the method 3 has possibility which it is actually effective.

We are planning several future works: (i) increasing the number of subjects and additional evaluation in the method 3, (ii) evaluation of actual ranking change.

References:

[1] A. Sato, J. Ogawa, and H. Kitakami: An Impression-based Retrieval System of Music Collection, *Proceeding of the 4th International Conference on KES2000*, Vol.2, 2000, pp.856-859.

[2] T. Hochin, K. Yamada, and T. Tsuji: Multimedia Data Access Based on the Sensitivity Factors, *Proceedings of the 2000 International Database Engineering and Applications Symposium(IDEAS'00)*, 2000, pp.319-326.

[3] T. Ikezoe, Y. Kajikawa, and Y. Nomura, Music Database Retrieval System with Sensitivity Words Using Music Sensitivity Space, *IPSJ(Information Processing Society of Japan) Journal*, Vol.42, No.12, 200, pp. 3201-3212.

[4] G. Salton, and M. J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, 1983.

[5] J. G. Snider, and C. E. Osgood, *Semantic Differential Technique – A Sourcebook*, Aldine Pub. Company, 1969.

[6] T. Bozkaya, and M. Ozsoyoglu: Indexing Large Metric Spaces for Similarity Search Queries, *ACM Transactions on Database Systems*, Vol.24, No.3, 1999, pp.361-404.

[7] T. Houchin, Personal Adaptation Method in the System Retrieving Multimedia Data Based on Impression, *Proceeding of the 9th World Multiconference on Systemics, Cybernetics and Informatics*, 2005, (CD-ROM).

[8] T. Takayama, T. Ikeda, Y. Takeda, and S. Kuroda: Proposition of Direct Interface for Multimedia Database Retrieval by the Combination of Impression Values, *Proceedings of the 2003 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, 2003, pp.609-612.

[9] B. Kurniawan: Java *Web Development with Servlets, Jsp, and Ejb 2nd Edition*, Sams Published, 2004.

[10] Oracle 9i Database: http://otn.oracle.co.jp/products/oracle9i/index.html