Received 15 January 2016
Revised 12 March 2016
Accepted 3 April 2016

Towards global music digital libraries

A cross-cultural comparison on the mood of Chinese music

Xiao Hu
University of Hong Kong, Hong Kong, Hong Kong, and
Jin Ha Lee
University of Washington, Seattle, Washington, USA

Abstract

Purpose – The purpose of this paper is to compare music mood perceptions of people with diverse cultural backgrounds when they interact with Chinese music. It also discusses how the results can inform the design of global music digital libraries (MDL).

Design/methodology/approach – An online survey was designed based on the Music Information Retrieval Evaluation eXchange (MIREX) five-cluster mood model, to solicit mood perceptions of listeners in Hong Kong and the USA on a diverse set of Chinese music. Statistical analysis was applied to compare responses from the two user groups, with consideration of different music types and characteristics of listeners. Listeners' textual responses were also analyzed with content coding.

Findings – Listeners from the two cultural groups made different mood judgments on all but one type of Chinese music. Hong Kong listeners reached higher levels of agreement on mood judgments than their US counterparts. Gender, age and familiarity with the songs were related to listeners' mood judgment to some extent.

Practical implications – The MIREX five-cluster model may not be sufficient for representing the mood of Chinese music. Refinements are suggested. MDL are recommended to differentiate tags given by users from different cultural groups, and to differentiate music types when classifying or recommending Chinese music by mood.

Originality/value – It is the first study on cross-cultural access to Chinese music in MDL. Methods and the refined mood model can be applied to cross-cultural access to other music types and information objects.

Keywords Digital libraries, Cross-cultural, Chinese music, Mood perception, Music digital libraries, Music mood

Paper type Research paper

1. Introduction

Music seeking and consumption are no longer confined by the boundaries of country, region or culture today (Lee *et al.*, 2013). Music, as a cultural object, may be perceived differently by people from different cultural backgrounds, imposing a challenge on music digital libraries (MDL) in meeting the needs of a diverse audience (Weissenberger, 2015). Consequently, an increasing number of researchers have started to investigate various cross-cultural issues in the music information retrieval (MIR) and MDL fields. As people often seek music for emotional goals (Lavranos *et al.*, 2015), music mood[1] has increasingly become a popular access point for music information in many MDL and online music services (Hu, 2010; Hu and Downie, 2007). This trend has raised questions regarding the applicability of music mood across cultural boundaries. Probably due to its subjective and context-based nature, music mood perception is often regarded as culture-dependent (Wong *et al.*, 2009). A number of previous studies have compared

Journal of Documentation
Vol. 72 No. 5, 2016
pp. 858-877

0022-0418
DOI 10.1108/JD-01-2016-0005

mood perceptions of music in various cultures by listeners from different cultural backgrounds (e.g. Balkwill and Thompson, 1999; Fritz *et al.*, 2009; Hu and Lee, 2012; Lee *et al.*, 2013; Singhi and Brown, 2014; Egermann *et al.*, 2014). Although specific findings vary, a general trend found in these existing studies is that listeners' perceptions of music mood can be influenced by their cultural backgrounds.

However, most existing studies have focussed on western (classical or pop) music or ethnic music with small audiences (e.g. Congolese music produced and consumed by Congolese Pygmies (Egermann *et al.*, 2014)). Such studies mainly aim to answer questions on music psychology and/or the social, developmental aspects of music. The purpose of this study is different; we are interested in exploring how music mood could be used as an access point or metadata type that can transcend the cultural boundaries of users in the context of MDL. To date, the question of how people with different cultural backgrounds would perceive the mood of Chinese music remains under-investigated. Considering its appeal to one of the largest music listener populations in the world (IFPI, 2014), Chinese music should not be left out of today's global scope of MDL/MIR studies. As Chinese music possesses unique characteristics that are remarkably different from western music (Fung, 2013), it is of theoretical and practical value to investigate the case of Chinese music.

Therefore, in this study, we aim to fill this gap and explore how people with different cultural backgrounds, specifically those in the USA and Hong Kong (representing western and eastern listeners, respectively), perceive the mood of Chinese music, including traditional Chinese music and contemporary Chinese pop music that infuses western pop characteristics and local Chinese elements. We specifically answer the following research questions:

RQ1. How differently or similarly do Hong Kong and US music listeners perceive the mood of Chinese music?

RQ2. How differently or similarly do Hong Kong and US music listeners perceive the mood of different genres/styles of Chinese music?

RQ3. Do the listeners' demographic characteristics and familiarity with the music affect their mood perceptions of Chinese music?

Answers to these questions have important implications for MDL that use mood as an access point to facilitate MIR and discovery in a global environment. These answers will help determine whether it is necessary to develop different mood-related labels or tags that take their suitability for various user groups into consideration. Implications of the results in this study will be discussed. The goal of this study is to contribute to the literature of cross-cultural MDL/MIR and provide empirical evidence for designing cross-cultural MDL/MIR systems.

2. Background and related work

2.1 A brief history and types of Chinese music

Chinese music broadly refers to music that uses traditional Chinese instruments or is in the Chinese language, and is the main music consumed in the region of Greater China (i.e. regions where Chinese is the official language, including mainland China, Hong Kong, Macau, Taiwan and Singapore). However, the consumption of Chinese music is not limited to that region. People around of the world are accessing Chinese music, and all major online music services, such as Spotify and Last.fm, contain Chinese music to cater to the demand.

There is a broad range of genres and styles in Chinese music. Traditional Chinese folk music dates back to approximately 7,000 years ago, played mostly in the pentatonic scale with its own set of instruments. Contemporary Chinese pop music (also referred to as C-pop) was developed in the middle of the twentieth century with the influence of western music. The subgenre "Cantopop" (Hong Kong pop music sung in Cantonese) was developed in the 1970s, followed by the rise of another subgenre, "Mandopop" (Mandarin Chinese pop music), in the 1980s (Fung, 2013). This study includes six genres/styles of Chinese music that are widely appreciated in the region and around the world. The first two are traditional Chinese music, whereas the rest are the most popular C-pop styles featured in online music services in China (e.g. Baidu, Xiami, Kugou):

- Traditional Instrumental Folk Music of China (*qiyue*): on the pentatonic scale, musicians perform using traditional instrument(s) of woodwind and percussion, bowed strings, plucked and struck strings in solo, small ensembles or orchestra (Randel, 2003).
- Canto(nese) folk songs: Han Chinese folk songs are divided into northern and southern styles geographically. Provinces south of the Yangzi River have mild weather and abundant rain, and therefore the folk songs tend to be more "lyrical, gentle and conjunct" (Han, 1989). Near the Canton province and Pearl River Delta, Cantonese folk music is especially light-spirited (China Culture, 2006).
- Contemporary Chinese instrumental music: this style of music is also called "New Music of China." It mixes the techniques, instrumentation and composition of both Chinese traditional music and the European/western music (Liu and Mason, 2010). The Girls 12 is a famous mainland Chinese all-female group who plays this type of music with traditional Chinese instruments.
- Cantopop (or Cantonese popular music): this is a subgenre of C-pop. Sung in Cantonese, Cantopop usually refers to popular songs produced in Hong Kong. Blended with traditional Chinese elements, as well as western-influenced harmonies and melodies, Cantopop was a dominant music genre in Asia in the 1980s' and 1990s. Its iconic star, Anita Mui, was recognized as "Asia's Madonna."
- China wind (*zhongguofeng*): this is a subgenre of Mandopop and is also called "Chinese Style." It features traditional Chinese instruments and classical Chinese melody juxtaposed with popular western music styles and cultural references to Chinese history, classics and folklore in the lyrics (Xinhuanet, 2006). Taiwanese singer Jay Chou is regarded as the "godfather" of this type of music, bringing forth the wave of China Wind pop.
- Chinese rock: as another subgenre of Mandopop, Chinese rock is locally produced in mainland China and is the most notable sub-category of Mandopop to the world outside of China. Chinese rock combines the loud and forceful folk song traditions of Northern China and characteristics of western-style rock such as a fast tempo, a strong beat and aggressive bass lines (Fung, 2013). The singer/songwriter Cui Jian from mainland China has been widely recognized as the father of Chinese rock (Fung, 2013).

It is noteworthy that all C-pop styles are, to varying extents, influenced by western pop or Rock music. Compared to Traditional Instrumental Chinese Folk Music and Canto Folk Songs, C-pop may be easier to understand and appreciate for listeners familiar with western music. In addition, because of the mixture of western music elements,

listeners from both cultural groups in this study might share more common perceptions on C-pop than the other two genres/styles. Therefore, in addition to analyzing aggregated perceptions across all songs, this study also compares and contrasts listeners' perceptions on these six music styles.

2.2 Cross-cultural studies on music mood perceptions

A number of studies have investigated the effects of culture on music mood perception. Although fundamental emotions such as happiness, sadness, anger or fear were found to be generally agreed upon across cultures, cultural traditions and background were found to affect individuals' music mood perception, in addition to the inherent qualities of the music (e.g. Gregory and Varney, 1996; Kosta *et al.*, 2013; Lee *et al.*, 2013). For example, Wong *et al.* (2009) compared mood descriptors used by European and Indian communities and discovered that cultural tradition played a more crucial role than the inherent qualities of the music in users' mood perception. Fritz *et al.* (2009) found that mood perceptions of western music by African listeners partially agreed with those by western listeners. In a similar vein, Hu and Lee (2012) discovered that discrepancies existed between Chinese and American music listeners in the way they determined the mood of western music pieces. More recently, Lee and Hu (2014) also compared perceptions of western music mood among American, Korean and Chinese listeners, and found that the results of Korean listeners' mood perceptions were situated in between the two other groups.

There are also studies investigating the cultural backgrounds of listeners in relation to their mood perceptions of non-western music. Kosta *et al.* (2013) found greater agreement on the mood of Greek songs among Greek listeners compared to non-Greek listeners with varying backgrounds. Similar results were found by Hu *et al.* (2014) in comparing American and Korean listeners' mood perceptions of a set of Korean pop songs. Egermann *et al.* (2014) compared subjective and psychophysiological emotional responses by Canadians and Congolese Pygmies to both western and Congolese music. They found that the two groups had very different responses to Pygmies music but similar responses to some western music.

To the best of our knowledge, there has been little work focussed on comparing the mood perceptions of various genres/styles of Chinese music across different cultural groups. In addition, many of the previous studies were conducted in the context of music psychology or ethnomusicology. This study, however, reflects the perspective of information organization and access, with the goal of finding out how music mood can be organized to best meet the information needs of users from different cultural groups.

3. Mood representation models in MIR

There are mainly two kinds of models to represent mood in music psychology and MIR: categorical and dimensional. The former uses a set of discrete terms (e.g. "passionate," "cheerful") to represent the mood of a piece of music. The most classical model of this kind is Hevner's (1936) adjective circle where eight mood categories placed in a circle, with a set of terms in each category. The dimensional models, in contrast, represent mood with continuous values in a low dimensional space. Different models may have different dimensions, yet valence (i.e. level of pleasure) and arousal (i.e. level of energy) are among the most popular dimensions. The dimensional model used very often in MIR is Russell's (1980) model. The categorical and dimensional models have their own pros and cons. Yang and Chen (2012) summarize that categorical models are more

user-friendly as they only consist of terms in natural language; whereas dimensional models are advantageous in terms of quantifying the intensity of moods. Notwithstanding the importance of dimensional models, categorical ones are more suitable for this current study, for the purpose of validating and comparing the values of mood metadata in the context of cross-cultural MDLs.

The models from music psychology have solid theoretical ground and can be borrowed into MIR studies. However, one criticism of adopting these models directly in MIR is that they were developed in experimental settings and thus might not reflect users' preferences and behaviors in today's music listening environment (Hu, 2010). Therefore, MIR researchers developed mood models based on domain analysis of widely used music repositories and streaming services.

One of the influential mood models in MIR is the five-cluster model proposed in the audio music mood classification task (Hu *et al.*, 2008) of Music Information Retrieval Evaluation eXchange (MIREX), a community-wide MIR evaluation campaign held annually since 2005 (Downie *et al.*, 2014). This model was developed from more than 180 editorial mood labels on AllMusic.com, a very popular online music repository. It has been used in many music mood classification studies (e.g. Laurier *et al.*, 2008; Bischoff *et al.*, 2009; Yang and Chen, 2012) as well as cross-cultural user studies (Hu and Lee, 2012; Lee and Hu, 2014; Hu *et al.*, 2014; Yang and Hu, 2012). The five mood clusters are:

- C1: passionate, rousing, confident, boisterous, rowdy;
- C2: rollicking, cheerful, fun, sweet, amiable/good natured;
- C3: literate, poignant, wistful, bittersweet, autumnal, brooding;
- C4: humorous, silly, campy, quirky, whimsical, witty, wry; and
- C5: aggressive, fiery, tense/anxious, intense, volatile, visceral.

In this five-cluster model, each cluster is collectively represented by a set of mood terms that share similar semantics. As the terms are describing affect, which can be somewhat vague (as opposed to terms referring to concrete objects), it is argued that using multiple terms with similar meanings to collectively describe one kind of mood is clearer than single-word representations (Hu and Downie, 2007).

A more recent categorical mood model was also used in MIREX, in the mood subtask of the audio tag classification task (Hu *et al.*, 2009). It consists of 18 mood categories derived computationally from mood-related social tags from Last.fm, arguably the largest music tagging and sharing website. This 18-class model consists of 135 tags. Besides the automated tag classification task in MIREX, a simplified version of this model has been used in a cross-cultural user study on Korean music (Hu *et al.*, 2014), together with the five-cluster model.

As the MIREX five-cluster mood model has been used in a number of cross-cultural MIR studies, for comparison purposes, the current study adopts this model and attempts to make suggestions on modifications and refinement based on the results.

4. Research design

4.1 The participant groups

The user groups recruited in this study consist of music listeners from Hong Kong and the USA. Hong Kong plays a special role in the history of Chinese music. Located adjacent to the Canton province, Hong Kong was the origin of Cantopop, which highly influenced the entire region of East Asia in the 1980s and 1990s. Furthermore, for

historical and political reasons, Hong Kong was able to preserve the traditional Chinese culture to a large extent. Canto folk music and even traditional Canto opera are often appreciated and practiced by Hongkongers. Since the government handover in 1997, people in Hong Kong also have had increasing exposure to Mandopop. Therefore, it makes an interesting case to investigate Hongkongers' perceptions of a variety of Chinese music types, since they are culturally exposed to all of them to certain degrees. On the other hand, US listeners can represent people with the western cultural background, and, additionally, they have been found to exhibit different music consumption behaviors than Hongkongers in a prior study (Nettamo *et al.*, 2006). We would like to explore whether differences also exist in the two user groups' mood perceptions of non-western music. Invitations were distributed to a number of departmental student mailing lists at the University of Hong Kong and University of Washington, as well as via the authors' various social network websites, such as Twitter, Facebook and Google Plus. Participation was completely voluntary.

4.2 Song selection

This study used an online questionnaire of 29 clips of Chinese music to solicit listeners' music mood perceptions. The clips were manually selected from the authors' personal music collections, with the goal of reflecting the broad range of Chinese music. As the goal of this study is to propose suggestions for global MIR systems, two kinds of Chinese music were considered: traditional Chinese music and contemporary Chinese pop music. The former features distinct cultural origin and musical characteristics, and thus can be attractive to users who do not have a Chinese background but are curious about Chinese or Asian music. The latter (C-pop) is widely consumed by Chinese-speaking listeners around the world. As it is heavily influenced by western pop music, yet has Chinese flavor, it is potentially of interest to users in all regions who like pop music in general. Specifically, the selected songs represented the following six genres or styles described in Section 2.1: traditional Chinese instrumental music, traditional Cantonese (Canto) folk music, contemporary Chinese instrumental music, Cantopop and two unique styles of Mandopop: China wind and Chinese rock. Each type contains five representative pieces except for Chine wind which contains four. The list of titles can be found at: http://ccmir.cite.hku.hk/pubs/moodCH/List_of_Songs.xlsx

4.3 The questionnaire

At the beginning of the questionnaire, participants were asked about their gender, year of birth and the place where they grew up. The last question was used for screening purposes so as to ensure that only participants with Hong Kong and US cultural backgrounds were included. We believe that people are greatly affected by the culture of the place they resided when they grew up, regardless of their nationality. All participants were 18 years of age or above and agreed with the informed consent statement before proceeding to the questionnaire.

Next, participants were asked to listen to the selected music clips, with one song on each page. The questionnaire consisted of five questions for each clip. The first question asked the participant to listen to the clip and decide the most appropriate mood expressed by the song. The response options given were the mood clusters of the MIREX five-cluster model. In case the participants found that the mood of the song did not fit into any of the five default clusters, they could select the option "Other," and use their own words to describe the appropriate mood of the music clip.

The participants were also asked about their familiarity with the music clips. Because people may interpret "familiarity" differently, instead of directly asking how familiar they are with the given songs, we used two questions: whether they had heard the song before, and if they could provide the artist name and the title of the song. Responses to the two questions were deemed more objective than self-rated familiarity because the questions are fact-based.

The questions were presented in English in the US questionnaire and in both English and Chinese in the Hong Kong questionnaire. In the Hong Kong questionnaire, the five mood clusters were also translated into Chinese, and both English and Chinese terms were presented. We used the same translation of mood terms as in Yang and Hu (2012). Even though English is one of the official languages in Hong Kong, people's English proficiency varies. Therefore, both Chinese translations and original terms in English were deemed necessary for the respondents to have an accurate interpretation of the clusters. Figure 1 illustrates the survey interface of the Hong Kong version.

4.4 Data analysis

Responses to the first question for each song (i.e. mood judgment) were compared quantitatively. The χ^2 test was used for comparing the distributions of music mood judgments between the two groups of listeners (Hinkle *et al.*, 2003). Significant results would indicate that the judgment distributions are dependent on cultural groups. The agreement between groups of listeners in various combinations of music and listener characteristics was measured by the raw ratio of agreed pairs to all possible pairs. For example, for calculating the agreement ratio among 32 Hong Kong listeners on five Cantopop songs, we first calculated the total number of judgment pairs as

Figure 1.
Screenshot of the survey questions for one song

Note: With both English and the Chinese translation

$(32 \times (32-1)/2) \times 5 = 2{,}480$, and counted the number of pairs with the same mood judgments. Suppose there were 248 matched pairs; the agreement ratio would be 10 percent. Raw agreement measures are recommended for cases where the annotations are dominated by negative values (i.e. "not chosen") (Cicchetti and Feinstein, 1990). This applies to the mood judgment data collected in this study. For instance, "Other" was chosen only 102 times out of 1,856 times it could have been selected.

Text responses to the "Other" option were coded with a scheme of categories that emerged from the analysis process. The codebook was developed through an iterative process involving test-coding and revising the codes for clarity. The consensus model (Hill *et al.*, 1997) was applied to coding the responses: two independent coders coded the responses and then discussed on disagreed cases until a consensus was reached. The results were compared across the two groups. Responses to demographic questions and song familiarity questions were used to group the users with a range of criteria, for the purpose of testing factors that affect users' mood perceptions.

4.5 Limitations

As each music piece needs to have a minimum length (typically 15-30 seconds) for listeners to establish a confident perception of its mood (Speck *et al.*, 2011), the questionnaire can contain only a limited number of song clips so that it can be completed within a reasonable amount of time. We used a set of 29 clips in this study, which may seem rather limited given there are millions of Chinese songs available in the market. To mitigate this limitation, we tried our best to carefully select a good variety of songs based on demands for them by users with various cultural backgrounds, and representativeness in each category. In addition, the sample size of each cultural group (32) is limited, partially due to the length of the survey and the lack of monetary incentive. Nonetheless, it is sufficient for statistical tests and comparisons. The online venues used to recruit participants (i.e. mailing lists and social network sites) may also have limited the participants to frequent internet users. Future studies are warranted to further verify the evidence found in this exploratory study in order to generalize findings to broader contexts.

5. Results

5.1 Characteristics of participants

The Hong Kong survey had a total of 70 responses, among which 34 were completed. The US survey had 72 responses, of which 34 were completed. However, two listeners in each group answered that they were raised in mainland China (not Hong Kong or the USA), leaving 32 valid responses for each group. Table I shows the demographic information of the listeners.

5.2 Mood judgments

5.2.1 Cultural backgrounds and mood judgments. Table II presents the distribution of mood judgments of the two groups on all 29 songs. Both groups chose Cluster 3

	Age				Gender	
Group	Min	Max	Average	SD	Female	Male
HK	24	60	36.1	9.30	12	20
USA	21	49	32.3	7.90	24	8

Table I. Demographic information of the listeners

(literate, poignant, wistful, etc.) most often, indicating that many Chinese songs may have some traits associated with the mood of Cluster 3, and these traits can be perceived by both Hong Kong and US listeners. The χ^2 test shows that the distribution and cultural group are not independent ($\chi^2 = 52.54$, df = 5, $p < 0.001$). In other words, it provides strong evidence that Hong Kong and US listeners tend to perceive the mood of given Chinese music differently, within the framework of five mood clusters. From Table II, we can observe that Hong Kong listeners tended to select Cluster 1 (passionate, rousing, confident, etc.) and Cluster 3 while US listeners tended to spread their choices equally across Clusters 1, 2 (rollicking, cheerful, fun, etc.) and 3. The US listeners were also more likely to select the "Other" category than their Hong Kong counterparts. In all, 26 out of the 29 songs were labeled with "Other" at least once by the US group. In comparison, the Hong Kong group chose the "Other" option for 17 songs. On average, 2.93 US listeners chose "Other" for each of these songs, whereas the number for Hong Kong listeners was 1.35. This is similar to the findings in Lee and Hu (2014), where US listeners chose the "Other" category more often than Korean and Chinese listeners when they made mood judgments on western songs. Even though the objects/stimuli have been changed to Chinese music, the observation still holds. The consistent pattern regardless of the cultural background of the music further suggests that this phenomenon might be related to the individualism-collectivism difference between western and eastern cultures as discussed in previous studies (Boer and Fischer, 2010; Hu and Lee, 2012; Lee and Hu, 2014): western users tend to hold their opinions regardless of the given options, whereas Eastern people are more likely to conform to existing norms.

We conducted a content analysis on the free-text responses provided in the "Other" option. The two independent coders reached an excellent reliability level (Cohen's $\kappa = 0.858$). Table III shows the seven resultant categories and corresponding counts in each of the groups. The largest category of the self-supplied terms is new mood terms that are not included in the five-cluster scheme. The new mood terms provided by both groups are quite similar, with, "relaxing," "calm" and "longing" being the most frequent, appearing six, four and four times, respectively. Users in both cultural groups also filled in terms that were not mood related. Some of them were music specific (e.g. "fast pace") while others were not (e.g. "focus").

Interestingly, the US group also provided responses that were either a mixture of terms from multiple clusters (e.g. "passionate" from Cluster 1 and "intense" from Cluster 5) or part of the terms in one cluster (e.g. "poignant, wistful, autumnal" in Cluster 3). The US participants also filled in descriptions of scenes they imagined when hearing the songs, although the instructions specifically asked for mood-related terms. Perhaps this group of users thought that scenes inspired by the songs also reflected the mood to some extent. The Hong Kong listeners did not provide any terms in these categories, possibly because: they tried harder to follow the instructions that asked

Table II.
Distribution of mood judgments

Group	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Other	Total
HK	257 (28%)	175 (19%)	289 (31%)	108 (12%)	76 (8%)	23 (2%)	928
USA	203 (22%)	199 (21%)	233 (25%)	100 (11%)	114 (12%)	79 (9%)	928

Note: Percentages are shown in parentheses

Categories	USA	HK	Examples
New mood terms	35	11	Calm; relaxing; courageous; 喜慶的(celebrational); 優閒的 (relaxing)
Description of scenes	13	0	Petals on the surface drifting with the wind; It makes me think of nature – water, leaves, sunrises; being powdered before dancing for the emperor;
Mixture of terms from different clusters	12	0	Passionate, intense; rollicking, cheerful, fun, whimsical, wry
Musical descriptions	7	5	Fast pace; 塞外風情 (Beyond the Great Wall style); traditional instrumentation with modern twist
Part of terms in a cluster	8	0	Rollicking, rousing, confident; poignant, wistful, autumnal
Non-mood, non-musical terms	3	5	Innovative; focus; 囉嗦的(verbose); familiar
No opinion	1	2	I have trouble selecting anything 情感起伏不大 (few emotional ups and downs)
Total	79	23	

Table III. Categories of terms provided in "Other" option

respondents to "choose one dominant cluster when a song sounds to have moods in more than one cluster," and people from individualistic (mostly western) cultures were found to be more used to divergent thinking which involves the generation of novel ideas (Schmidt *et al.*, 2013).

5.2.2 Agreement on mood judgments within and across culture groups. The agreement ratios were calculated for pairs of judgments between two listeners from the same cultural group and across cultural groups. The results are shown in Table IV. In particular, in the cross-cultural case, the number of total pairs with one judgment from each cultural group is calculated as $((32 \times 32)/2) \times 29 = 29{,}696$. The results and non-paired t tests reveal that the level of intra-agreement among Hong Kong listeners is significantly higher than that of US listeners ($p < 0.001$). A similar result was reported in Kosta *et al.* (2013), where Greek listeners were found to agree more on the mood of Greek music than non-Greek listeners. In comparing Korean and American listeners' perceptions of the mood of Korean pop songs, Hu *et al.* (2014) also found that Korean listeners reached a higher agreement level than their US counterparts. However, the result is partially different from Hu and Lee (2012), where both US and Chinese listeners had the same agreement ratio on the mood of western pop songs. We conjecture that this difference in results between Hu *et al.* (2014) and Hu and Lee (2012) might be due to Chinese listeners' wide exposure to western pop songs. For non-western songs that have not yet spread as widely, sharing the same cultural background as the music stimuli does seem to help listeners reach a higher agreement on music mood.

It is also noteworthy that the agreement level among US listeners was not statistically different from the cross-cultural agreement ($p = 0.54$). This is similar to the result in

Group	Number of agreed pairs	Number of total pairs	Ratio
HK	6,394	14,384	0.44
USA	5,491	14,384	0.38
Across	11,451	29,696	0.39

Table IV. Agreement on pairs of mood judgments

Hu *et al.* (2014), where the agreement level among American listeners on the mood of Korean pop songs was the same as that between American and Korean listeners.

5.2.3 Agreement on mood clusters. To examine which moods are easier or more difficult to agree upon, we counted the number and percentage of agreed pairs across the five mood clusters (Table V). The χ^2 test indicates that the two cultural groups showed a statistically significant difference in the agreement pattern across the mood clusters ($\chi^2 = 603.55$, df = 10, $p < 0.001$).

Both within and across cultures, Cluster 3 (literate, poignant, wistful, etc.) got the most agreements, which again evidences that many Chinese pop songs have traits of Cluster 3 and these traits can transcend the cultural boundaries of listeners to some extent. The difference between the two cultural groups observed here is similar to the findings in Hu and Lee (2012), where American listeners agreed more on Cluster 2 (rollicking, cheerful, fun, etc.), Cluster 4 (humorous, silly, campy, etc.) and Cluster 5 (aggressive, fiery, tense/anxious, etc.) compared to Chinese listeners, who agreed more on Cluster 1 (passionate, rousing, confident, etc.) and Cluster 3 (literate, poignant, wistful, etc.). As western music was used as the stimuli in Hu and Lee (2012), it is interesting to see that the patterns of mood perception are similar to this study even though the stimuli here consist of completely different, non-western music.

5.3 Mood perception on different types of music

5.3.1 Instrumental vs vocal. The music pieces in this study could be divided into instrumental and vocal sets. The former includes 15 pieces in instrumental folk, Canto folk or contemporary instrumental styles, and the latter includes the rest of the three types with 14 pieces total. We can observe from Table VI that overall, Hong Kong listeners agreed among themselves more often than US listeners on both instrumental and vocal pieces. Listeners also agreed more often on vocal pieces than instrumental music within each cultural group, as well as cross-culturally. This is in accordance with the literature on both music psychology and user studies in MDL/MIR. Juslin (2005) argued that the verbal (lyric) and non-verbal (melody) channels of music complemented

Group	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Other	Total
HK	1,962 (31%)	884 (14%)	2,719 (43%)	391 (6%)	430 (7%)	8 (0%)	6,394
USA	1,151 (21%)	1,107 (20%)	1,900 (35%)	550 (10%)	664 (12%)	119 (2%)	5,491
Across	2,911 (25%)	2,092 (18%)	4,681 (41%)	639 (6%)	1,058 (9%)	70 (1%)	11,451

Table V.
Agreement across mood clusters

Note: Number and percentage (in parentheses) of agreed pairs across five mood clusters among listeners within and across cultural groups

Group	Instrumental	Vocal	All
HK	0.38	*0.51*	0.44
USA	0.35	*0.41*	0.38
Across	0.34	*0.44*	0.39

Table VI.
Agreement ratio on instrumental and vocal music

Note: Agreement ratio on instrumental and vocal music, with the highest ratio in each group marked in italics

each other in communicating emotions. Vocal delivery and the lyrics of the song were also found to be among the various factors that influence people's perception of music mood (Lee *et al.*, 2012). In comparing the effects of melodies and lyrics on mood perceptions, studies found that, although melodies were more dominant than lyrics, lyrics were particularly useful in strengthening emotional responses to music with negative (Ali and Peynircioglu, 2006) and low-arousal emotions (Kazuma, 2009), which roughly correspond to Clusters 5 and 3, respectively, in this study. Even though the US listeners probably did not understand the lyrics in those songs, they might still have been able to get some emotional clues from the vocal delivery.

5.3.2 Genres/styles of Chinese music. Table VII shows the agreement ratios within and across cultural groups with regard to the six types of music. The Hong Kong listeners showed an agreement of 60 percent on Cantopop, followed by 54 percent on China wind. The US listeners agreed more on Canto folk and China wind (43 percent), followed by Cantopop (41 percent). Cross-culturally, χ^2 tests show that the two groups significantly differed in their mood judgments on all styles except for Cantopop ($\chi^2 = 9.22$, $\mathrm{df} = 10$, $p = 0.10$). We conjecture two possible reasons for Cantopop to reach a high cross-cultural agreement ratio: as introduced before, Cantopop was developed with heavy influence from western pop music, and thus the western elements could have helped to bridge the differences between the two user groups; and all Cantopop songs are lyrical and thus the mood cues in the vocal delivery may have helped strengthen mood perceptions. The category with the least agreement both within and across cultures was contemporary instrumental music. This may be due to the complex musical compositions of this style and the lack of lyrical cues. In fact, the song with the least agreement was of this type and will be analyzed in more detail in the next subsection.

To facilitate further examination of the mood judgments on different types of music stimuli, Figure 2 shows the distribution of mood judgments for each of the six music types included in this study. It can be seen that the first three types of music (i.e. instrumental folk, Canto folk and contemporary instrumental) were labeled with a mix of diverse mood clusters without a clear, dominant cluster, whereas the last 3 (i.e. Cantopop, China wind and Chinese rock) have one dominant cluster from both cultural groups: China wind and Cantopop were mostly labeled with Cluster 3 (literate, poignant, wistful, etc.) and Chinese rock with Cluster 1 (passionate, rousing, confident, etc.). Instrumental folk music seems to be the most controversial, in that the pieces got similar numbers of judgments across all mood clusters, especially from the US group. The two cultural groups seemed to disagree most on Canto folk and contemporary instrumental music. Hong Kong listeners categorized Canto folk music mostly in

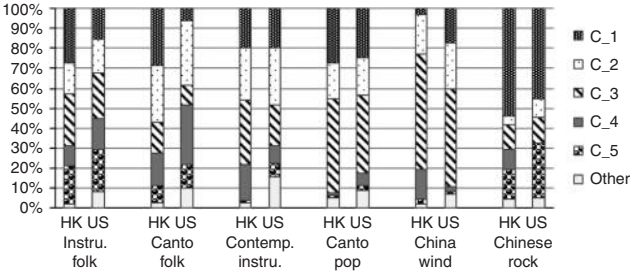

Figure 2.
Distribution of mood clusters in different types of Chinese music

Clusters and 2 (rollicking, cheerful, fun, etc.) but US listeners tended to label those pieces more often with Cluster 2 and Cluster 4 (humorous, silly, campy, etc.). For contemporary instrumental music, US listeners chose Cluster 2 the most, followed by similar numbers for Cluster 3, Cluster 1 and "Other," while Hong Kong listeners used Cluster 3 more often, followed by Cluster 2, Cluster 1 and Cluster 4.

5.3.3 The songs with highest and lowest agreement. The song that reached the highest agreement on mood judgment was the same within each cultural group and across cultural groups. In fact, the US listeners agreed more than Hong Kong listeners, with agreement ratios of 0.94 and 0.88, respectively. This was a China wind song titled *Pearl-decorated Curtain Rolled Up* by a Chinese male singer, Huozun. In all, 30 Hong Kong listeners and 31 US listeners selected Cluster 3 (literate, poignant, wistful, etc.) for this song. Released in 2014, the song features a slow melody and extensive use of falsettos. It can be characterized as mellow, ethereal and graceful, similar to the style of ancient Chinese music (and thus recognized as China wind). Although the lyrics describe a beautiful young woman missing a lost love in a poetic way, the US listeners probably could not understand them. Yet they reached near consensus on the mood of this song, possibly attributable to the emotionally powerful delivery of the song.

The song with the lowest agreement among US listeners as well as cross-culturally was Shangri-la by Girls 12, a female band famous for using traditional Chinese instruments to play both traditional Chinese and western music. This song had the second-lowest agreement among Hong Kong listeners, and their lowest agreement was for another piece by the same band. Many pieces by Girls 12 are modern compositions using ancient instruments, including Shangri-la, and thus are classified as contemporary instrumental music, the type that has the least agreement both within and across cultures (Table VII). The song Shangri-la received only 16 percent agreement from US listeners and 21 percent from Hong Kong listeners. Figure 3 illustrates the distribution of mood judgments of this song. We can observe that the judgments from Hong Kong listeners were distributed fairly evenly across Cluster 1

Table VII.
Agreement ratio across different music types, with the highest ratio in each group marked in italics

Genre/style	No. of songs	HK	US	Across
Instrumental folk	5	0.41	0.33	0.36
Canto folk	5	0.37	*0.43*	0.37
Contemporary instrumental	5	0.36	0.30	0.32
Cantopop	5	*0.60*	0.41	*0.49*
China wind	4	0.54	*0.43*	0.44
Chinese rock	5	0.41	0.39	0.38

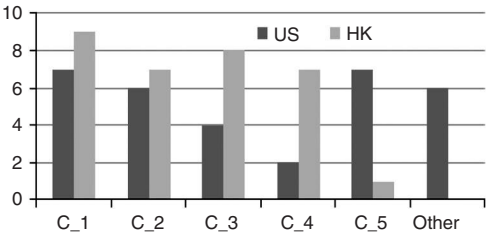

Figure 3.
Comparison of mood judgments of Hong Kong and US listeners on Shangri-la

(passionate, rousing, confident, etc.), Cluster 2 (rollicking, cheerful, fun, etc.), Cluster 3 (literate, poignant, wistful, etc.) and Cluster 4 (humorous, silly, campy, etc.), and about equal numbers of US listeners labeled it with Cluster 1, Cluster 2, Cluster 5 (aggressive, fiery, tense/anxious, intense, etc.) and "Other." The piece was performed with traditional Chinese instruments: bamboo flute, *erhu* (bowed strings), *pipa* and *guzheng* (both plucked strings), mixed with western-sounding percussion and electronica fusions. Just as the song title suggests, the piece conjures the mystic and mysterious aura of the utopia named Shangri-la. The mixed mood perceptions of this song can be at least partially explained by the mixed musical nature of the song, as well as its fictional, mystical theme.

5.4 Listener characteristics

5.4.1 Gender. Table VIII shows the distribution of mood judgments based on gender across all mood clusters. Both genders in both cultural groups selected Cluster 3 the most. The χ^2 test indicates that gender difference in mood judgment does exist within the US group ($\chi^2 = 24.62$, df = 5, $p < 0.001$), but not in the Hong Kong group ($\chi^2 = 7.53$, df = 5, $p < 0.184$). This was also the case in Lee and Hu (2014), where American listeners' mood judgments on western music differed based on gender, but Chinese listeners' did not.

In order to see the influence of gender difference across cultural groups, we compared the judgments from US and Hong Kong listeners of the same gender. The comparison shows female listeners in the two cultural groups did have different mood judgments ($\chi^2 = 46.51$, df = 5, $p < 0.001$). However, the judgments of male listeners from the two cultural groups did not have a significant difference ($\chi^2 = 7.99$, df = 5, $p = 0.157$). The results suggest that cultural differences played a significant role in mood judgments among female listeners, but not among male listeners. In contrast, Hu and Lee (2012) found significant differences for both genders between US and Chinese listeners when the music stimuli were western pop songs. This discrepancy might have been affected by the small sample size of male participants in the US group ($n = 8$) in this study. Future studies with a larger and more balanced sample across gender groups are warranted to further investigate the extent to which gender and cultural groups affect music mood perception.

5.4.2 Age. The ages of the participants vary between the two groups, but both groups have listeners between 21 and 60 years old. We divided those listeners into two age groups, 21-33 and 34-60, so as to balance the sample sizes in both age groups. The χ^2 tests on the distribution of judgments revealed a statistically significant difference

Group	HK		USA	
Gender	Female	Male	Female	Male
C1	104	153	147	56
C2	61	114	147	52
C3	*112*	*177*	*165*	*68*
C4	35	73	73	27
C5	32	44	87	27
Other	4	19	77	2
Total	348	580	696	232

Table VIII. Mood judgments across genders and mood clusters, with the highest number in each group marked in italics

between age groups among Hong Kong listeners ($\chi^2 = 39.86$, df = 5, $p < 0.001$) but no significant difference among US listeners. ($\chi^2 = 4.72$, df = 5, $p = 0.45$). This is, in fact, consistent with the findings from Lee and Hu (2014), where the authors found a significant difference between the age groups for Korean and Chinese music listeners but not for American listeners. This could be possibly attributed to the fast-changing environment in Hong Kong and other Asian countries compared to the relatively stable environment in countries such as the USA (Lee and Hu, 2014).

5.4.3 Familiarity with the music. Listeners' familiarity with the songs was categorized into three levels: "high familiarity" if they had listened to the song before and could also identify the artist's name and the song title; "medium familiarity" if they had heard the song but could not name it; and no familiarity if they had not heard the song. There was one US listener who was highly familiar with two songs, and seven who, collectively, had heard a total of 15 songs before. In the Hong Kong group, 29 out 32 listeners indicated high familiarity with at least one song. Unsurprisingly, more than half of the highly familiar songs were Cantopop.

Due to the sparseness of data in the US group, we examined how the levels of agreement related to familiarity with the music only in the Hong Kong group. Table IX lists the agreement ratios among judgment pairs with regard to various levels of music familiarity. When both participants were highly familiar with a song, their agreement was the highest (ratio = 0.57). While a higher familiarity level helped increase the level of agreement when both mood judgments were made for songs with which participants were familiar, it did not help when one participant was familiar with the song and the other was not. We conjecture that when users have heard a song multiple times or know the artist name and the song title, they may have developed and/or gained some contextual knowledge of the song, which may affect their selection of mood clusters. On the other hand, when listeners hear the song for the first time, they may naturally rely more on the musical characteristics for judging the mood of the song.

6. Implications for cross-cultural MDL/MIR design

As advocated by Weissenberger (2015), a flexible organization system is needed for MIR/MDL to meet the needs of different musical traditions. The results of this study have important implications for designing MDL of this kind.

As users in different cultural groups had significantly different mood judgments overall (Section 5.2.1), it is better to differentiate user-generated mood labels and/or mood-related social tags according to the users' cultural background. In particular, as the agreement ratio among Hong Kong listeners was higher than the cross-cultural one (Section 5.2.2, Table IV), Hong Kong listeners would benefit more from mood tags supplied by listeners in the same cultural group. In contrast, as the within-group agreement ratio of US listeners was similar to the cross-cultural agreement level (0.38 vs 0.39), US listeners could make use of mood tags generated by Hong Kong listeners, especially when those generated by US listeners are sparse.

Table IX. Agreement ratio across different levels of music familiarity

Familiarity	No	Medium	High	*n*
No	0.43	0.37	0.42	549
Medium		0.43	0.52	202
High			0.57	177

The analysis on the textual input in the "Other" option (Section 5.2.1) suggests that the MIREX five-cluster model may not be sufficient and more terms are needed to describe the mood of Chinese music. Particularly, terms like "calm," "relaxing," "peaceful" and "longing" appeared most frequently (Table III, Section 5.2.1). In fact, a previous study on automated mood classification on Chinese pop songs (Yang and Hu, 2012) also used mood terms in the MIREX five-cluster model, and ended up with adding additional mood labels that were deemed as important to represent moods in Chinese pop songs: "calm/peaceful," "dreamy," "encouraging," "nostalgic," "relaxed," "soothing" and "tender." Based on this finding, it is suggested a sixth cluster can be added to the MIREX five-cluster model: (calm, relaxing, peaceful, etc.). In addition, some of the suggested terms in the current study were mixtures of terms from different clusters, which implies that the mood clusters may not be mutually exclusive, at least for US listeners. Lee and Hu (2014) also found that some songs may be perceived to be in two or more clusters at the same time. Therefore, it is suggested that the MIREX five-cluster model should allow multiple mood clusters to be applied to each song.

As US listeners specified terms that describe visual scenes, it might be a good idea to use scenes as a new access point in MDL. The association between music and images has been studied in music psychology (Ali and Peynircioglu, 2006) and MIR (Mattek and Casey, 2011), but how it can be used in MDL warrants future research to explore.

The fact that both cultural groups chose Cluster 3 (literate, poignant, wistful, etc.) most frequently (Table II) and agreed more on songs labeled with Cluster 3 (Table V) suggests that Chinese music may tend to have musical traits that are associated with moods in Cluster 3, and that such musical traits can transcend the cultural boundaries of the audience. MDL using automated mood classification algorithms (Yang and Chen, 2012) could exploit this finding and assign more weight to the classifier of Cluster 3, allowing songs to have a greater chance of being classified in Cluster 3. Similarly, when recommending songs based on mood labels, those labeled with Cluster 3 could be assigned more confidence.

As different styles had different agreement levels within and across cultures (Table VII), music in different styles can be treated differently in MDL. For example, contemporary instrumental pieces had low agreement levels between audiences, and thus classification and recommendations of pieces in this style could be given less confidence or provide more candidates to improve the chance of hits.

The results on gender and age group differences can be exploited, too. For instance, Hong Kong listeners would benefit more from recommendations and tags made by other Hong Kong listeners in the same age group, whereas for US listeners there is no need to differentiate the age group of other listeners. The effect of song familiarity on mood perception suggests that MDL should allow users to change their mood labels and tags over time, because when they get more familiar with a song, they may perceive the mood of the song differently.

7. Conclusion and future work

This study compares Hong Kong and US listeners' mood perceptions of 29 Chinese music pieces, with the goal of investigating whether and how mood perceptions of the two groups of listeners differ, and how the differences can inform the design of cross-cultural MDL. Music mood was modeled with the MIREX five mood clusters and the results suggested further refinement of this model. The selected music pieces included six genres and styles of Chinese music, ranging from traditional folk music to several sub-genres of C-pop.

Overall, the patterns of mood judgments by listeners from Hong Kong and the USA (e.g. commonly agreed-upon clusters, differences based on gender and age) resonated with the findings in previous studies on cross-cultural mood perceptions of western music, even though this study used Chinese music, which has a contrasting cultural background to that of western music. This has important implications for designing global MDL: for both western and Chinese pop music, listeners from different cultures may need different mood labels. For systems that use social tags to facilitate access, it is recommended that tags given by users in different cultural groups be differentiated. Those given by users in the same cultural groups can be more helpful. This study also confirms that listeners do exhibit cross-cultural differences in mood judgment for all but one type (Cantopop) of Chinese music considered in this study. Therefore, it may be desirable for MDL to differentiate music types when classifying or recommending music by mood.

Future work will include further investigation on other cultural groups interested in Chinese music, such as mainland Chinese, Taiwanese and Singaporeans. In addition, we plan to conduct similar studies on music from other cultural backgrounds, such as Korean and Japanese music. Findings of such studies will further contribute to the goal of building effective and user-centered global MDL.

Note

1. In the literature, music mood is also referred to as music emotion or the affective aspect of music. In this paper we use "mood" for consistency.

References

Ali, S.O. and Peynircioglu, Z.F. (2006), "Songs and emotions: are lyrics and melodies equal partners?", *Psychology of Music*, Vol. 34 No. 4, pp. 511-534.

Balkwill, L.L. and Thompson, W.F. (1999), "A cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues", *Music Perception*, Vol. 17 No. 1, pp. 43-64.

Bischoff, K., Firan, C.S., Paiu, R., Nejdl, W., Laurier, C. and Sordo, M. (2009), "Music mood and theme classification – a hybrid approach", *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR), Kobe*, pp. 657-662.

Boer, D. and Fischer, R. (2010), "Towards a holistic model of functions of music listening across cultures: a culturally decentred qualitative approach", *Psychology of Music*, Vol. 40 No. 2, pp. 179-200.

China Culture (2006), "Canton music (in Chinese)", available at: www.chinaculture.org/gb/cn_whyc/2006-11/20/content_89367.htm (assessed September 8, 2015).

Cicchetti, D.V. and Feinstein, A.R. (1990), "High agreement but low kappa: II. Resolving the paradoxes", *Journal of Clinical Epidemiology*, Vol. 43 No. 6, pp. 551-558.

Downie, J.S., Hu, X., Lee, J.H., Choi, K., Cunningham, S.J., Hao, Y. and Bainbridge, D. (2014), "Ten years of MIREX (Music Information Retrieval Evaluation eXchange): reflections, challenges and opportunities", *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR), Taipei*, pp. 657-662.

Egermann, H., Fernando, N., Chuen, L. and McAdams, S. (2014), "Music induces universal emotion-related psychophysiological responses: comparing Canadian listeners to Congolese Pygmies", *Frontiers in Psychology*, Vol. 5, pp. 1341-1349.

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R. and Koelsch, S. (2009), "Universal recognition of three basic emotions in music", *Current Biology*, Vol. 19 No. 7, pp. 573-576.

Fung, A. (2013), "Deliberating fandom and the new wave of Chinese pop: a case study of Chris Li", *Popular Music*, Vol. 32 No. 1, pp. 79-89.

Gregory, A.H. and Varney, N. (1996), "Cross-cultural comparisons in the affective response to music", *Psychology of Music*, Vol. 24 No. 1, pp. 47-52.

Han, K.-H. (1989), "Folk songs of the han Chinese: characteristics and classifications", *Asian Music*, Vol. 20 No. 2, pp. 107-128.

Hevner, K. (1936), "Experimental studies of the elements of expression in music", *The American Journal of Psychology*, Vol. 48 No. 2, pp. 246-268.

Hill, C.E., Thompson, B.J. and Williams, E.N. (1997), "A guide to conducting consensual qualitative research", *The Counseling Psychologist*, Vol. 25 No. 4, pp. 517-572.

Hinkle, D.E., Wiersma, W. and Jurs, S.G. (2003), *Applied Statistics for the Behavioral Sciences*, Wadsworth, Cengage Learning, Belmont, CA.

Hu, X. (2010), "Music and mood: where theory and reality meet", *Proceedings of the 5th iConference*, University of Illinois at Urbana-Champaign, Champaign, IL.

Hu, X. and Downie, J.S. (2007), "Exploring mood metadata: relationships with genre, artist and usage metadata", *Proceedings of the 8th International Society for Music Information Retrieval Conference (ISMIR)*, *Vienna*, pp. 67-72.

Hu, X. and Lee, J.H. (2012), "A cross-cultural study of music mood perception between American and Chinese listeners", *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR), Porto*, pp. 535-540.

Hu, X., Downie, J.S. and Ehmann, A. (2009), "Lyric text mining in music mood classification", *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR)*, *Kobe*, pp. 411-416.

Hu, X., Lee, J.H., Choi, K. and Downie, J.S. (2014), "A cross-cultural study of mood in K-POP songs", *Proceedings of the 15th International Society for Music Information Retrieval Conference (ISMIR)*, *Taipei*, pp. 385-390.

Hu, X., Downie, J.S., Laurier, C., Bay, M. and Ehmann, A. (2008), "The 2007 mirex audio mood classification task: lessons learned", *Proceedings of the 9th International Society for Music Information Retrieval Conference (ISMIR), Philadelphia, PA*, pp. 462-467.

IFPI (2014), "The International Federation Of The Phonographic Industry: IFPI digital music", available at: www.ifpi.org/downloads/Digital-Music-Report-2014.pdf (accessed September 8, 2015).

Juslin, P.N. (2005), "From mimesis to catharsis: expression, perception, and induction of emotion in music", in Miell, D., MacDonald, R. and Hargreaves, D.J. (Eds), *Musical Communication*, Oxford University Press, Oxford, pp. 85-115.

Kazuma, M.O.R.I. (2009), "The influence of the meaning of lyrics on the expressed emotion of music valence", *Proceedings of the 2nd International Conference of Students of Systematic Musicology*, *Ghent, November 18-20*, pp. 53-58.

Kosta, K., Song, Y., Fazekas, G. and Sandler, M.B. (2013), "A study of cultural dependence of perceived mood in Greek music", *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR), Curitiba*, pp. 317-322.

Laurier, C., Grivolla, J. and Herrera, P. (2008), "Multimodal music mood classification using audio and lyrics", *7th IEEE International Conference on Machine Learning and Applications (ICMLA)*, *December*, pp. 688-693.

Lavranos, C., Kostagiolas, P.A., Martzoukou, K. and Papadatos, J. (2015), "Music information seeking behaviour as motivator for musical creativity: conceptual analysis and literature review", *Journal of Documentation*, Vol. 71 No. 5, pp. 1070-1093.

Lee, J.H. and Hu, X. (2012), "Generating ground truth for mood classification in music digital libraries using mechanical turk", *Proceedings of the IEEE-ACM Joint Conference on Digital Libraries*, pp. 129-138.

Lee, J.H. and Hu, X. (2014), "Cross-cultural similarities and differences in music mood perception", *iConference 2014 Proceedings*, pp. 259-269.

Lee, J.H., Hill, T. and Work, L. (2012), "What does music mood mean for real users?", *Proceedings of the 2012 iConference*, *February*, pp. 112-119.

Lee, J.H., Choi, K., Hu, X. and Downie, J.S. (2013), "K-pop genres: a cross-cultural exploration", *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR), Curitiba*, pp. 529-534.

Liu, J. and Mason, C. (2010), *A Critical History of New Music in China*, Chinese University Press, Hong Kong.

Mattek, A. and Casey, M. (2011), "Cross-modal aesthetics from a feature extraction perspective: a pilot study", *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR), Miami*, pp. 585-590.

Nettamo, E., Nirhamo, M. and Häkkilä, J. (2006), "A cross-cultural study of mobile music: retrieval, management and consumption", *Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*, pp. 87-94.

Randel, D.M. (2003), *The Harvard Dictionary of Music*, Vol. 16, Harvard University Press, Cambridge.

Russell, J.A. (1980), "A circumplex model of affect", *Journal of Psychology and Social Psychology*, Vol. 39 No. 6, pp. 1161-1178.

Schmidt, J.J., Facca, T. and Soper, J.C. (2013), "International variations in divergent creativity and the impact on teaching entrepreneurship", *Journal of Higher Education Theory and Practice*, Vol. 13 No. 2, pp. 101-109.

Singhi, A. and Brown, D.G. (2014), "On cultural, textual and experiential aspects of music mood", *Proceedings of the 15th International Society for Music Information Retrieval Conference*, pp. 3-8.

Speck, J.A., Schmidt, E.M., Morton, B.G. and Kim, Y.E. (2011), "A comparative study of collaborative vs traditional musical mood annotation", *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR), Miami*, pp. 549-554.

Weissenberger, L. (2015), "Toward a universal, meta-theoretical framework for music information classification and retrieval", *Journal of Documentation*, Vol. 71 No. 5, pp. 917-937.

Wong, P.C., Roy, A.K. and Margulis, E.H. (2009), "Bimusicalism: the implicit dual enculturation of cognitive and affective systems", *Music Perception*, Vol. 27 No. 2, pp. 81-88.

Xinhuanet (2006), "Sounds a lot like China - Hong Kong and Taiwan's music field stirs up 'Chinese style'", available at: http://news.xinhuanet.com/overseas/2006-10/11/content_5187968.htm (assessed December 28, 2015).

Yang, Y.H. and Chen, H.H. (2012), "Machine recognition of music emotion: a review", *ACM Transactions on Intelligent Systems and Technology (ACM TIST)*, Vol. 3 No. 3, pp. 40-69.

Yang, Y.H. and Hu, X. (2012), "Cross-cultural music mood classification: a comparison on English and Chinese songs", *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR), Porto*, pp. 19-24.

Further reading

Chow, Y.F. and de Kloet, J. (2010), "Blowing in the china wind: engagements with Chineseness in Hong Kong's zhongguofeng music videos", *Visual Anthropology*, Vol. 24 Nos 1-2, pp. 59-76.

Gwet, K.L. (2010), *Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Raters*, Advanced Analytics, LLC, Gaithersburg.

Rana, S.A. and North, A.C. (2007), "The role of music in everyday life among Pakistanis", *Music Perception*, Vol. 25 No. 1, pp. 59-73.

Corresponding author

Xiao Hu can be contacted at: xiaoxhu@hku.hk