Understanding Emotion in *Raag*: An Empirical Study of Listener Responses

Parag Chordia and Alex Rae

Georgia Institute of Technology,
Music Technology Group
840 McMillan St, Atlanta GA 30332, USA
{ppc,arae3}@gatech.edu
http://www.paragchordia.com

Abstract. A survey of emotion in North Indian classical music was undertaken to determine the type and consistency of emotional responses to *raag*. Participants listened to five one-minute *raag* excerpts and recorded their emotional responses after each. They were asked to describe the emotions each excerpt evoked and then to adjust six different sliders indicating the degree to which they felt the following: happy, sad, peaceful, tense, romantic, longing. A total of 280 responses were received. We find that both free-response and quantitative judgments of emotions are significantly different for each *raag* and quite consistent across listeners. We hypothesized that the primary predictors of emotion in these excerpts would be pitch-class distribution, pitch-class dyad entropy, overall sensory dissonance, and note density. Multiple regression analysis was used to determine the most important factors, their relative importance, and their total predictive value (R^2). The features in combination explained between 11% (peaceful) and 33% (happy) of response variance. For all models, a subset of the features were significant, with the interplay between "minor" and "major" scale degrees playing an important role. Although the explanatory power of the current models is limited, the results thus far strongly suggest that *raags* do consistently elicit specific emotions that are linked to musical properties. The responses did not differ significantly for enculturated and non-enculturated listeners, suggesting that musical rather than cultural factors are dominant.

1 Background

Raag literally means "that which colors the mind". It is a melodic abstraction around which almost all North Indian classical music (NICM) is organized. A *raag* is most easily explained as a collection of melodic gestures and a technique for developing them. The gestures are sequences of notes that are often inflected with various micro-pitch alterations and articulated with an expressive sense of timing. Longer phrases are built by joining these melodic atoms together.

NICM uses approximately one hundred *raags*, of which fifty are common. Despite micro-tonal variation, the tones in any given *raag* conform to a subset of

R. Kronland-Martinet, S. Ystad, and K. Jensen (Eds.): CMMR 2007, LNCS 4969, pp. 110–124, 2008.

the twelve chromatic pitches of a standard just-intoned scale. There are theoretically thousands of scale types; in practice, however, *raags* conform to a much smaller set of scales, and many of the most common *raags* share the same set of notes.

By building phrases as described above, a tonal hierarchy is created. Some tones appear more often in the basic phrases, or are sustained longer. Indian music theory has a rich vocabulary for describing the function of notes in this framework. The most stressed note is called the *vadi* and the second most stressed, traditionally a fifth or fourth away, is called the *samvadi.* There are also less commonly used terms for tones on which phrases begin and end. A typical summary of a *raag* includes its scale type (*that*), *vadi* and *samvadi.* A pitch-class distribution (PCD), which gives the relative frequency of each scale degree, neatly summarizes this information.

The performance context of *raag* music is essentially monophonic, although vocalists will usually be shadowed by an accompanying melody instrument. The rhythmic accompaniment of the *tabla* is also present in metered sections. There is usually an accompanying drone that sounds the tonic and fifth using a harmonically rich timbre.

2 Related Work

The emotional qualities of *raags* have been discussed extensively in NICM music theory. Traditionally, this music is said to evoke seven basic emotions: sadness, romance, peace, strength/courage, anger, dispassion, devotion [1]. In theory, each *raag* elicits a unique emotional state (*rasa*) consisting of one or more of these basic emotions.

For Western music, a variety of studies have been undertaken since the late 19^{th} century to define what emotions music elicits, and to understand the musical factors that underly them. These studies have followed a basic experimental paradigm in which musical excerpts are presented to listeners who are then asked to respond by verbally describing their emotional states, rating emotions on a quantitative scale, or some other measurement. Musical stimuli are chosen so that certain characteristics can be systematically varied, such as timbre, tempo and tonality. Correlations between responses and musical structure are then observed, and are interpreted to account for some aspect of the listener reactions.

A common pitfall in such experiments is the fundamental difficulty of isolating the factor that is being tested. Real musical excerpts rarely vary in just one parameter, while artificial stimuli that can be systematically varied often will not generalize, lacking the "ecological validity" of real music. Most importantly, the difficulty of quantifying a person's actual emotional state, either through observation or by eliciting verbal or other responses, makes such research extremely challenging. Despite this, the field has produced a body of work that suggests consistent relationships between musical structure and emotional response [2]. In general, two broad categories of emotional responses have been found, relating to valence (happy-sad) and activity (vigorous-calm). For example, fast tempo has been shown to correlate consistently with positive valence and activity; the same has been shown for major (versus minor) tonality. A thorough overview of the field of emotion and music can be found in [3].

Although the connection between *raag* and emotion is assumed to be systematic by nearly all enculturated listeners and performers, very little empirical work relating to emotion in NICM has been published. In one study, NICM musicians were asked to play short excerpts that conveyed either joy, sadness, anger, or peace. Listeners not familiar with NICM were asked to identify the predominant emotion. The recognition of intended emotions was mostly successful except for peace [2].

3 Motivation

The current work seeks to characterize the emotional responses of listeners to different *raags* and the extent to which listeners report similar experiences; furthermore, if responses differ systematically by *raag*, it seeks to identify some of the musical factors that underly these responses. In NICM, all pitch-related activity is structured by the choice of *raag*. Each *raag* is traditionally associated with several basic emotions. For example, *Raag Yaman* is said to evoke peace, happiness, and devotion. Performers and listeners of NICM often report that the essence of *raag*-based music is the evocation of mood brought about by one or more related emotions, and that different *raags* reliably elicit different emotional states. In many cases, however, the more traditional descriptions are not accepted, or are thought of as rudimentary. Despite the central role of emotion in *raag* theory, almost no empirical work has been undertaken to systematically analyze listener response to different *raags*.

4 Survey Design

For the purpose of collecting a large number of responses, we created an online survey in which participants were asked to respond to five *raag* excerpts in terms of "how listening to each *raag* made you feel." The survey can be found at `http://www.paragchordia.com/survey/raagemotion1/` and the *raag* excerpts can be heard at `http://www.paragchordia.com/research.html`. Some basic demographic information was collected, specifically age, sex, level of musical training (if any), and familiarity with NICM (if any). For each excerpt, there was both a free response with the directions to "describe how it makes you feel in as much detail as possible," and a series of six value sliders (ranging from zero to one hundred) corresponding to pre-chosen emotional adjectives: happy, sad, tense, peaceful, romantic, and longing. The adjectives were chosen based on preliminary research by the author suggesting that these might best capture the dominant emotional axes. Further, the first four roughly represent the valence and activity judgments discussed above.

Participants were asked to "adjust the slider to indicate how well you feel that the word applies to the *raag* excerpt you just heard." The order in which

Table 1. Summary of scale degrees used by *raags* in database. Notes are listed with C as the tonic.

	C	D♭	D	E♭	E	F	F♯	G	A♭	A	B♭	B
Yaman	•		•		•		•	•		•		•
Desh	•		•		•	•		•		•		•
Khamaj	•		•		•	•		•		•	•	•
Darbari	•		•	•		•		•	•		•	
Marwa	•	•			•		•			•		•

each of these sliders appeared was randomized for each *raag* presented. Each participant was presented with the same set of five *raags* in a random order. The specific excerpt for each *raag* was chosen at random from three possibilities, one played on *sarod* and two on *sitar*, both traditional plucked string instruments. All excerpts were approximately one minute long, unaccompanied, and from *alap*, a slower, arrhythmic introductory section of a *raag* performance. Participants had the option of replaying each segment an unlimited number of times. Without knowing in advance how much data we would be able to collect, it was impossible to determine an appropriate dimensionality for the data we sought. Therefore, an attempt was made to strike a balance, leading to the decision to limit the number of slider-values to six and include a free-response section as well. This was also the reasoning that determined on the one hand the presentation of the same five *raags* to each participant, while generalizing slightly by the inclusion of three audio files for each *raag*.

5 Survey Results

A total of 280 survey responses were collected. Participants were recruited from online NICM forums, as well as through the Georgia Institute of Technology music department. 29% of the respondents were female and 71% male, with a median age of 24. The respondents described their familiarity with NICM as "none" (11%), "a little"(24%), "somewhat"(29%), "very"(32%) and "expert"(4%).

5.1 Content Analysis of Free Responses

The descriptive language of the free responses indicated the intensity and complexity of listeners' emotions. The style ranged from lists of words and phrases describing simple emotions to lengthy and even poetic passages. A sampling is shown in Table 2.

Due to the fact that many of the free responses consisted of fragments or lists of words, highly sophisticated content analysis was not particularly relevant; instead, we focus on lower-level analysis, primarily word histograms.

Initially, simple histograms of terms appearing in each free response section were tabulated. Limiting the results to relevant descriptive terms, the most common for each *raag* were collected and are shown in Table 3. Standard variations of each word (e.g. "happy" and "happiness") were treated as examples of the same root word.

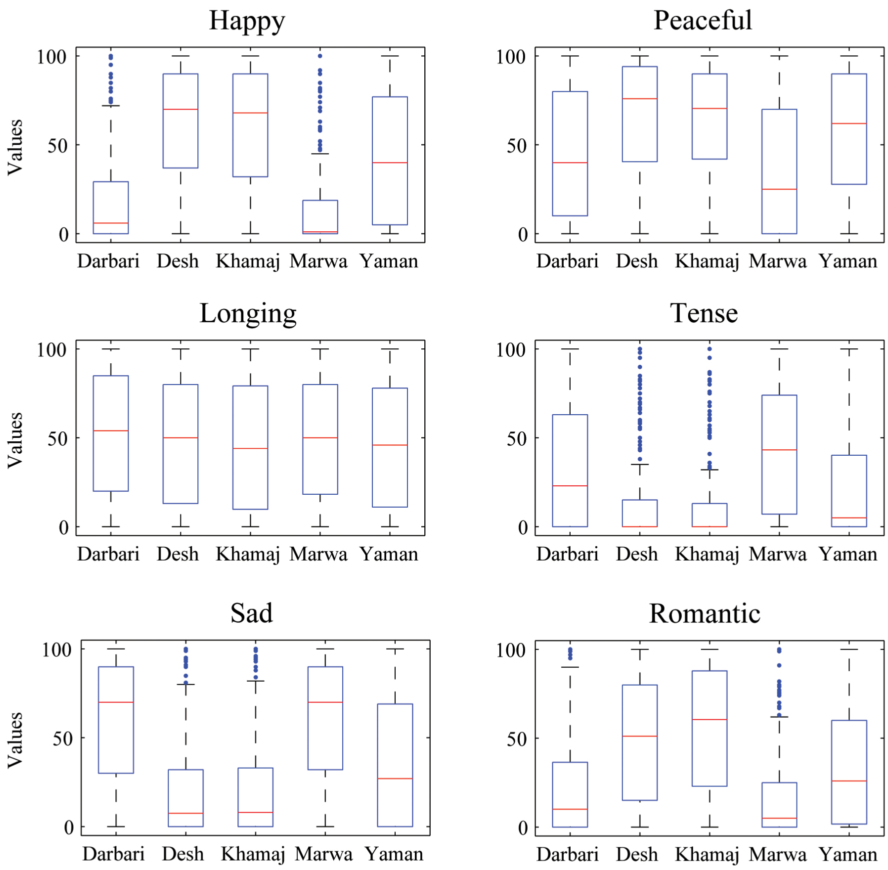

Fig. 1. Summary of quantitative survey responses. Box and whisker plots for each emotion are shown. The 25^{th} to 75^{th} percentile is indicated by the box, the median by the horizontal line in the middle of the box. The range of the data is shown by the whiskers with outliers notated with hatch marks. It can be seen that *Desh* and *Khamaj* form one reponse cluster, and *Marwa* and *Darbari* another.

Terms were also grouped according to various categories of meaning, in an attempt to both give a broader view of the responses and to capture information buried in the "long tail" of infrequently used but clearly related terms. For example, words such as "bliss", "exultation", "happy", "joyful", and "ecstasy" were grouped together under "Joy & Happiness". Relevant semantic groupings were based on a modified version of a standard content analysis dictionary. As before, the most common occurrences were tabulated and the results are displayed in Table 4.

The tables clearly show a common pattern of responses for certain raags. *Raags Khamaj* and *Desh* are associated with positively valenced concepts such as "joy & happiness," whereas *Darbari* and *Marwa* were consistently associated with "melancholy & sadness." Interestingly, *Yaman* was equally associated with

Table 2. A sample of free responses

Darbari: "Emptiness. I'm tumbling down a deep abyss. Weightless, then maybe. Dark and surrounding. Fall."
"It feels like pain, an agony that is long lasting and is happening at the time that the music is playing, a time of hardship."

Desh: "Absolutely fresh.... Clearing all the bad thoughts... Gives fresh meaning to life. Sounds serene. Very Peaceful i felt."
"Pure, unblemished...Something white as milk, offering to take you in, clean all your sins...Compassionate and loving, but in a distant sort of way."

Khamaj: "It so reminds me of the blossoming of flowers and prosperity... the happy chuckles of newborns... the pride of their mothers... "
"spring time. birds chirping. sun is shining. love. a mother is telling her child a story. the child is smiling. good times are upon us."

Marwa: "my reaction to this raaga was almost sexual. I felt desire. I felt the urge to look beautiful and dress up in heavy gold jewelery. felt very aware of my body. at times I felt a strange anger and the desire to control somebody else. i felt very powerful. i felt like a woman."
"very depressing...felt like crying...someone was going away... describes life in some way....the ups and downs."

Yaman: "I feel like a butterfly. The wind is streaking past me, and colors are awash in the air. Melodious colors."
"A combination of moods. It looks relaxed most of the time and ruminating about something. It seems to turn a little angry occasionally as though an unhappy event was inadvertently recollected."

Table 3. Simple word histograms for each *Raag*

Darbari		Desh		Khamaj		Marwa		Yaman	
Freq.	Term	Freq.	Term	Freq.	Term	Freq.	Term	Freq.	Term
75	sad	59	happy	53	happy	74	sad	46	sad
19	peaceful	41	peaceful	36	peaceful	16	serious	42	peaceful
14	deep	19	romantic	31	sad	14	good	38	happy
13	calm	17	sad	24	romantic	14	peaceful	13	good
10	longing	15	soothing	13	soothing	14	longing	12	calm
10	serious	11	nice	10	good	13	deep	11	longing
10	beautiful	10	pleasant	9	longing	10	tense	8	serious
9	good	10	patriotic	9	love	7	emotional	7	romantic
8	happy	9	playful	9	relax	7	happy	7	relax
5	sombre	9	remind	6	calm	6	soothe	6	devotional

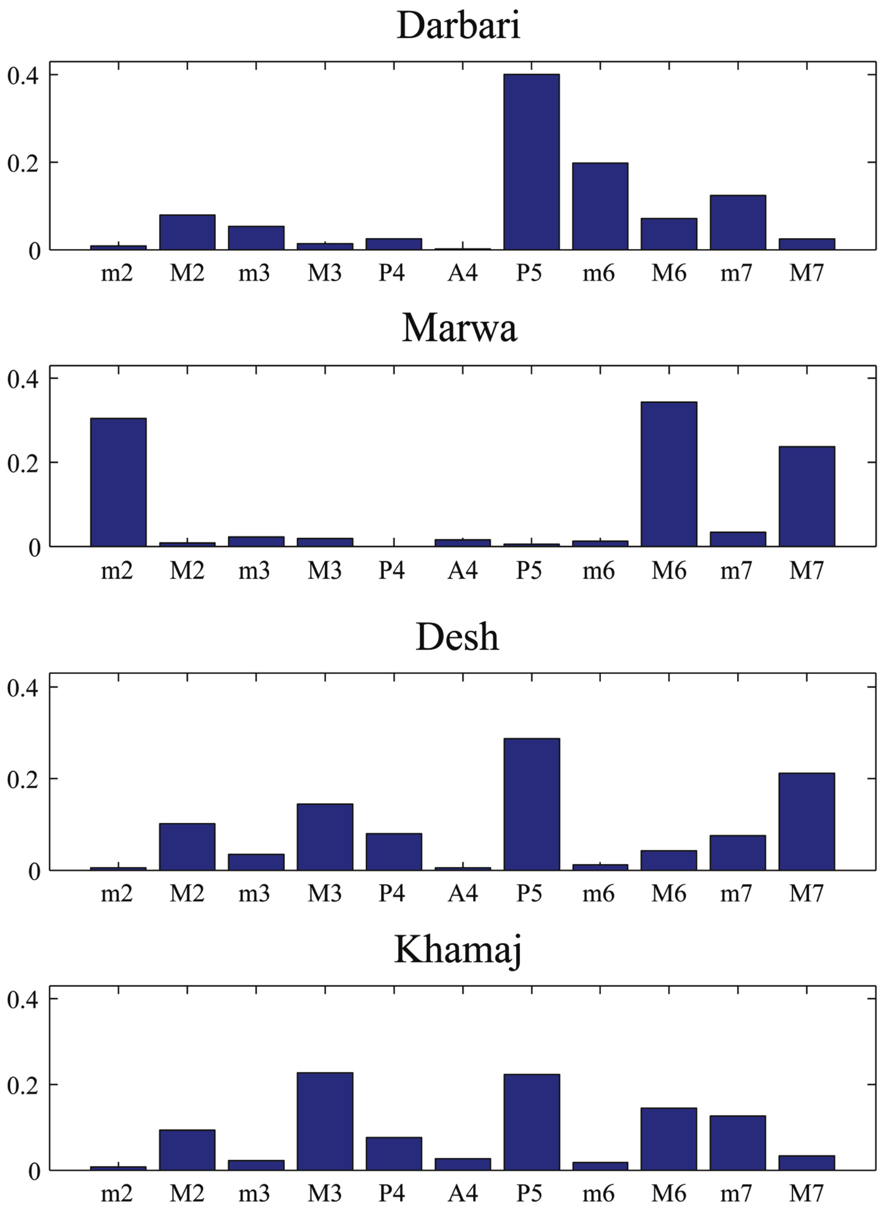

Fig. 2. Pitch-class distributions for *raags Darbari*, *Marwa*, *Desh*, and *Khamaj*

"melancholy & sadness" and "joy & happiness." The results show that, although each *raag* has a strong valence, a variety of other emotions are also present, suggesting that listeners' responses are indeed complex.

Table 4. Semantic category histograms for each *Raag*. Words with a similar meaning were assigned to a common semantic category.

Darbari		Desh		Khamaj	
Frq.	Semantic Cat.	Frq.	Semantic Cat.	Frq.	Semantic Cat.
134	melancholy & sadness	115	joy & happiness	103	joy & happiness
21	joy & happiness	41	love & affection	54	melancholy & sadness
16	pains	32	melancholy & sadness	37	love & affection
15	emotion	14	pleasures	13	emotions
12	beauty	13	composure	10	composure
11	tension	12	beauty	9	pleasures
11	love & affection	11	hope & optimism	7	passions
10	sympathy & compassion	6	emotion	7	hope & optimism
8	despair & resignation	6	desires	5	desires

Marwa		Yaman	
Frq.	Semantic Cat.	Frq.	Semantic Cat.
133	melancholy & sadness	71	melancholy & sadness
28	tension	69	joy & happiness
15	sympathy & compassion	16	emotions
12	pains	16	love & affection
12	joy & happiness	12	beauty
11	emotion	9	composure
10	love & affection	8	tension
7	anger & indignation	8	pleasures
7	emotionalities	6	passions

5.2 Quantitative Responses

Figure 2 shows the distribution of quantitative (slider) responses for each emotion and *raag*. For happy, sad, tense, and romantic, three *raag* clusters are apparent. *Desh* and *Khamaj* are strongly positively valenced (happy, romantic and not tense) and *Marwa* and *Darbari* are negatively valenced (sad, not romantic, tense) with *Yaman* falling between these poles. Multiple comparison of means confirms that differences between these three clusters for these three emotions are significant ($p < .05$) but that differences within clusters are not significant. The "longing" values show no statistically significant differences by *raag*, however all the *raags* clearly brought out this emotion. For the "peaceful" emotion there are again three clusters: *Marwa* is least peaceful followed by *Darbari* and then a cluster containing *Yaman*, *Khamaj* and *Desh*.

6 Feature Selection

In order to explore possible predictors for these emotional reactions, several features were extracted from the audio excerpts. Some of these features have previously been shown to be highly effective in *raag* classification [4], thus we sought to investigate their ability to explain emotional responses. We hypothesized that survey responses would in part be predicted by the pitches used and their relative weights, as represented by PCDs. In particular, it was expected that the relative prevalence of certain major and minor intervals would correspond to positively and negatively valenced responses.

Further, it was thought that sequential pairs of pitches, pitch class dyad distributions (PCDDs), would have a noticeable correlation. Extrapolating from Huron [5], we hypothesized that the degree of flexibility in PCDDs would correspond to a sense of tension and longing. Flexibility was calculated as the entropy of the PCDD distributions, where a low entropy would correspond to an excerpt having highly determined, or leading, note transitions, possibly leading to a greater sense of longing.

A number of other features were considered as well, notably the overall dissonance of a given excerpt, calculated as the sensory dissonance. This feature was thought likely to correspond to negative emotions, although as calculated it might give excessively large values to spectrally rich sounds. The overall note density and spectral centroids for each excerpt were also determined.

6.1 Pitch Detection and Pitch-Class Distributions

Pitch detection was done using a version of the Harmonic Product Spectrum (HPS) algorithm [6]. Each segment was divided into 40 ms frames, using a Gaussian window. The frames were overlapped by 75%, so that the pitch was estimated every 10 ms.

The PCDs were calculated by taking histograms of the pitch-tracks. The bins corresponded to each note of five octaves of a chromatic scale. Specifically, the ratios of the just-intoned scale tuned to the tonic frequency of each segment were used to calculate the bin values, and the bin boundaries were determined as the log mean. The five octaves were then folded into one, transforming pitches into pitch-classes, and the values were normalized to create a pitch-class distribution. This nulled any significance of octave errors, a common problem in the HPS algorithm.

6.2 Note Density

The average note density for each excerpt was calculated from detected note onsets. Onsets were found by thresholding a complex detection function (DF) [7]. The excerpt was divided into 128 sample regions, overlapped 50% using a rectangular window, and the DFT of each region was computed and used to construct the DF. Adaptive thresholding, proportional to the median over a sliding window, was used to choose the peaks to be labeled as onsets.

6.3 Pitch-Class Dyad Distributions

Pitch-class dyad distributions were calculated for each excerpt. The detected onsets were used to segment the pitch-tracks into notes. Each note was then assigned a pitch-class label: first the raw pitch estimates were discretized by assigning to each the center value of the bins defined for the pitch histogram, and then the mode was calculated for each note. The label of the corresponding chromatic pitch was assigned to that note. This process dealt quite effectively with variations due to micro-pitch structure, attacks, and errors by the detection algorithm. The octaves were folded into one as with the PCDs. The pitch-classes were then arranged in groups of two (bi-grams), or in musical terms, dyads. The entropy of the PCDD for each excerpt was then calculated.

6.4 Sensory Dissonance

The presence in a sound of partials which fall within the critical band for frequency resolution is the key indicator of dissonance, according to the tonotopic theory of sensory dissonance. A value for this was calculated by performing pairwise comparisons between all detected partials in each segment, and weighting the detected sub-critical-band intervals by their relative amplitudes. This calculation was performed using an algorithm developed by Kameoka and Kuriyagawa [8].

6.5 Timbral Features

The center of mass in the frequency domain (spectral centroid) was calculated for each excerpt.

7 Analysis

An initial demographic analysis showed that responses were not significantly different by age, sex, or familiarity with Indian music.

Multiple regression analysis was undertaken to determine how well the PCDs, subjective dissonance, and sensory dissonance features explained the emotion ratings. A linear model was built using these features for each of the quantitative emotional response types (with the exception of "longing", which was excluded because no statistically significant differences were found between *raags*). For example, the slider values for "happy" were modeled as

$$happy_i = \beta_1 x_1^i + \beta_2 x_2^i + \beta_3 x_3^i + \; + \beta_n x_n^i, \tag{1}$$

where x_k^i is the k$^{\text{th}}$ feature and β_k^i is the corresponding regression coefficient for the i$^{\text{th}}$ observation. Stepwise regression was performed to determine the subset of features to be included in the final model.

Tables 5, 6, 7, and 8 show the *beta* values, that is the standardized regression coefficients, for the each of the features, as well as the R^2 statistic for the total model. All coefficients presented have p values of less than .01. All variables in the model were normalized to have a mean of zero and a standard deviation of one, thus avoiding difficulties in comparing the scales of the independent variables. The R^2 statistic measures the total variance accounted for by the model and varies between 0 (random) and 1 (perfect fit). Of course, if measurement error is high, then the model will not be meaningful. The individual correlation between each feature and the emotion, which varies between -1 and 1, is also shown.

Additionally the variance inflation factor (VIF) for each feature is shown, giving a measure of the multicollinearity of the independent variables. This occurs when one of the variables is well approximated by a linear combination of the other variables, as is likely to be the case for PCD values. The VIF is calculated by regressing each variable using the remaining independent variables, and is computed as $\frac{1}{1-R^2}$. If the R^2 value is high, it indicates that the variable is multicollinear. In typical applications, VIF values above a threshold of 10 indicate that a variable should be removed. Interpretation of models with multicollinear variables is difficult as the relative contribution of each cannot easily be read from the beta values and the reliability of the estimate decreases (the size of the confidence interval increases for that coefficient). In some cases, multicollinearity can reverse the expected sign of the regression coefficient. For example, in Table 5 in the "happy" model, the Major 3^{rd} coefficient is negative despite a strong positive correlation; it can be seen, however, that the variable is highly multicollinear, with a VIF of 7.74. Because scale tones in *raags* may be highly correlated, it is important to be aware of this potential confound. Stepwise regression partially avoids this problem by automatically selecting features that give the greatest incremental contribution to explaining the variance of the dependent variable. In this way, only one of a set of highly correlated features is typically included. Nevertheless, highly correlated variables may remain in the model, making VIF useful for interpretation.

Tables 5 and 6 shows that "happy" and "sad" are modeled best, with R^2 values of .34 and .28 respectively, whereas the model is only weakly predictive in the case of "peaceful" (.11) and "tense" (.16). Unsurprisingly, "happy" responses are negative correlated with the minor 2^{nd} ($\beta = -.13$), minor 3^{rd} ($\beta = -.16$),

Table 5. Summary of regression model for "happy". For each feature we report the standardized regression coefficients, the bivariate correlation, and the variance inflation factor measure of multicollinearity.

Total $R^2 = .34$

feature	β	**corr**	**VIF**
minor 2^{nd}	-0.13	-0.29	4.23
minor 3^{rd}	-0.16	-0.09	1.86
Major 3^{rd}	-0.13	0.38	7.74
Perfect 4^{th}	0.13	0.37	5.66
minor 6^{th}	-0.43	-0.25	6.02
Major 6^{th}	-0.35	-0.28	2.33
Major 7^{th}	-0.37	-0.03	8.11
sensory dissonance	0.13	0.06	2.74
PCDD entropy	0.18	0.38	2.27

Table 6. Summary of regression model for "sad"

Total $R^2 = .28$

feature	β	**corr**	**VIF**
minor 2nd	0.14	0.26	2.11
Major 2nd	-0.16	-0.14	1.91
minor 3rd	0.35	0.09	2.12
Perfect 4th	-0.20	-0.29	1.66
minor 6th	0.20	0.28	1.36
Major 6th	0.20	0.25	2.29
spectral centroid	-0.18	-0.10	1.98
note density	-0.15	-0.11	1.65

Table 7. Summary of regression model for "peaceful"

Total $R^2 = .11$

scale deg	β	**corr**	**VIF**
Major 2nd	0.22	0.11	3.32
minor 3nd	-0.24	-0.05	2.76
Major 3rd	0.22	0.22	4.69
Perfect 4th	0.15	0.23	4.63
Perfect 5th	0.15	0.10	1.20
Major 7th	0.23	0.03	1.88
PCDD entropy	-0.16	0.19	3.46

Table 8. Summary of regression model for "tense"

Total $R^2 = .16$

scale deg	β	**corr**	**VIF**
minor 2nd	0.10	0.25	1.92
Major 2nd	-0.21	-0.15	1.71
minor 3rd	0.24	0.08	1.55
Major 3rd	-0.14	-0.26	1.45
Perfect 5th	-0.20	-0.18	1.74
minor 6th	0.14	0.12	1.45

minor 6th ($\beta = -.43$), and minor 7th ($\beta = -.37$). Surprisingly, the Major 3rd is also negatively correlated with a β value of -.13. As mentioned before, however, the Major 3rd has a high VIF value; examination of the correlation coefficients shows that it is strongly correlated with the Perfect 4th (.84), which is positively related to "happy" responses. "Sad" responses show an opposite

Table 9. Comparison of three regression models. The total adjusted R^2 value for each model is shown.

	Full	PCD only	Minor
Happy	0.34	0.27	0.07
Sad	0.28	0.22	0.06
Peaceful	0.10	0.10	0.02
Tense	0.16	0.13	0.05
Romantic	0.21	0.18	0.08

pattern, positively related to the minor 2^{nd} ($\beta = .14$), minor 3^{rd} ($\beta = .35$), and minor 6^{th} ($\beta = .20$), and negatively to the Perfect 4^{th} ($\beta = -.20$). Many of the other models loosely conform to the idea that *raags* with "minor" intervals are negatively valenced (sad, tense) and *raags* with "major" intervals are positively valenced (happy, peaceful).

The entropy of the PCDD distribution was a significant factor in the "happy" and "peaceful" models, positively related in the former and negatively related in the latter. We had hypothesized that low PCDD entropy values would correspond to tension and longing and higher values to a greater sense of flexibility. The effect shown here is weakly consistent with this prediction.

It is important to note that the relationship between PCDD entropy and the emotional characteristics we are testing is likely non-linear. Values outside the range represented in this study might well be expected to elicit different reactions. One might expect very low values, corresponding to predictability and repetition, to elicit "peaceful" and "sad" reactions. Very high values, on the other hand, might correspond to unpredictability and hence elicit feelings of "tension" and "stress". However, in the middle of the range, where most of the musical excerpts used here lie, low entropy corresponds to *raags* with a relatively fixed phraseology, creating a sense of tendency rather than repetition, and high entropy corresponds to *raags* with a greater degree of flexibility, conveying more variability than instability. If correlations observed in the above models are valid, this is the most likely explanation.

Two additional models were developed to see if a more parsimonious explanation of the data could be given. The first used the total strength of the minor 2^{nd} , minor 3^{rd} , and minor 6^{th} in the *raag* as the independent variable, a measure of the total degree of "minorness". The adjusted R^2,

$$R_a^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1}, \tag{2}$$

which allows comparison between models with a different number of independent variables, is reported in Table 9. The second model considered only the PCD features. The full model was significantly more explanatory. Although the full model suggested that a feature that combined the minor 2^{nd}, minor 3^{rd} and minor 6^{th} in total measure of "minorness" might capture most of the information in the PCD, this was not the case. The PCD features explained an additional 10-20% of variation as compared with the single "minor" feature.

8 Discussion

Survey responses have shown that different *raags* evoke a clearly differentiated set of emotional reactions. Free responses tended to cluster strongly in particular adjectival categories based on *raag*, and quantitative responses were significantly different by *raag* for all emotions except "longing". Thus, a substantial step has been taken towards empirical verification of the nature and reliability of emotional responses to *raag*. Importantly, responses did not vary systematically by familiarity with NICM suggesting that listeners were not simply referring to culturally determined concepts, but responding to underlying features of the music.

The analysis, although preliminary, suggests that responses are in part attributable to pitch-class statistics; the prevalence of certain scale degrees is useful in predicting the valence of the emotional responses. The data suggest that the entropy of the PCDD and the spectral centroid are also important. These are undoubtedly just a few of the many factors that influence listener responses. As more data are collected it will be possible to more fully examine other factors.

It is important to note that these models are currently merely suggestive. In none of the cases were they highly predictive, with a maximum of 34% of the variance accounted for. Because the goal here was explanatory rather than classificatory, the models were not verified on an independent data set. As with any task that forces respondents to verbalize primarily non-verbal mental states, there is significant measurement error due to the inherent unnaturalness of the task and an imperfect ability to map the verbal space. It is also possible that much of the true emotional feel of the music is lost in the projection onto simple emotions such as "happy" and "sad". Although it is likely that some aspect of *raags* can be effectively captured by mapping onto these axes, it is also likely that it is a gross simplification of the actual emotional experience.

9 Conclusions and Future Work

We have reported the results of the first empirical survey of listeners' emotional reactions to *raag* music. We have established that responses exhibit clear patterns and have identified several musical features which partially explain them. Based on the current work, we are in the process of conducting a more detailed and expansive survey. This will consider a larger set of *raags* and provide listeners with more dimensions along which to evaluate their emotional experiences. As more data is collected, it will be possible to test a greater number of musical features and more robustly assess their validity.

References

1. Bhatkande, V.: Hindusthani Sangeet Paddhati. Sangeet Karyalaya (1934)
2. Balkwill, L.L., Thompson, W.F.: A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues. Music Perception 17, 43–64 (1999)
3. Juslin, P.N., Sloboda, J.A.: Music and Emotion: Theory and Research. Oxford University Press, Oxford (2001)
4. Chordia, P., Rae, A.: Raag recognition using pitch-class and pitch-class dyad distributions. In: Proceedings of International Conference on Music Information Retrieval (2007)
5. Huron, D.: Sweet Anticipation: Music and the Psychology of Expectation. MIT Press, Cambridge (2006)
6. Sun, X.: A pitch determination algorithm based on subharmonic-to-harmonic ratio. In: Proc. of International Conference of Speech and Language Processing (2000)
7. Duxbury, C., Bello, J.P., Davies, M., Sandler, M.: A combined phase and amplitude based approach to onset detection for audio segmentation. In: Proc. of the 4th European Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2003), London, pp. 275–280 (2003)
8. Kameoka, A., Kuriyagawa, M.: Consonance theory, part i: Consonance of dyads. Journal of the Acoustical Society of America 45(6), 1451–1459 (1969)