Analyzing Music to Music Perceptual Contagion of Emotion using a Novel Contagion Interface: A Case Study of Hindustani Classical Music

Sanga Chaki[1], Sourangshu Bhattacharya[2], Raju Mullick[3], and Priyadarshi Patnaik[3]

[1] Advanced Technology Development Centre, IIT Kharagpur
[2] Computer Science and Engineering Department, IIT Kharagpur
[3] Department of Humanities and Social Sciences, IIT Kharagpur

Abstract. Music has strong potential to convey and elicit emotions, which are dependent on both context and antecedent stimuli. However, there is little research available on the impact of antecedent musical stimuli on emotion perception in consequent musical pieces, when one listens to a sequence of music clips with insignificant time lag. This work attempts to understand how the perception of one music clip is affected by the perception of its antecedent clip, with special reference to Hindustani Classical Music. We call this phenomenon perceptual contagion of emotion in music. Findings suggest, when happy clips are preceded by sad and calm clips, perceived happiness increases. When sad clips are preceded by happy and calm clips, perceived sadness increases. Calm clips are perceived as happy and sad when preceded by happy clips and sad clips respectively. This suggests that antecedent musical stimuli have capacity to influence the perception of music that follows.

Keywords: Music Perception, Perceptual Contagion of Emotion, Web based Self Report Surveys, Hindustani Classical Music (HCM)

1 Introduction

Emotions are an integral part of life. Under different circumstances, we feel happy, excited, angry, afraid, sad and many other emotions. It has been established by researchers [1], [2], that various stimuli and situations like social functions [3], interpersonal encounters [4], exposure to visual and auditory stimuli can elicit emotional responses. Music is one such potent stimuli [5]. In this discussion, we focus on the perceptive and cognitive aspects of music listening, which form effective ways of expressing and inducing emotions [6]. The perception of emotion in the consequent musical clip might be influenced by the perception of its antecedent clip – much like the meaning of sentences are inferred with reference to the preceding words or phrases. We performed web-based surveys to track if and how such influences occur. It was found that the perception of emotional content of music is modified by preceding music and its emotional content, even when the exposure is for a short duration. Thus, emotion perception in music is a sequentially modified and complex process.

The rest of section 1 documents the previous works done on emotion perception, how human emotions are captured for various studies and how emotions are represented in Indian philosophy. In section 2, we describe in details the procedure followed for the surveys and the stimuli used for the same. Section 3 shows the results of the context-based biases obtained in perception of music through tables and line charts. Conclusions are drawn and future scope is discussed in section 4.

1.1 Emotion Perception and Induction

There is a distinction between the perception and induction of musical emotions. In the first case, we only become aware of an emotion and in the second, we actually feel those emotions. The underlying processes might be different leading to requirement of apt representations and measurement models to study these phenomena. Juslin [6] explained how music arouses emotions by a multi-level framework, which included the phenomenon of *Contagion* among many other psychological mechanisms. In these studies by Juslin [6], [7], *Emotional contagion* refers to a process whereby an emotion is induced by a piece of music because the listener perceives the emotional expression of the music, and then mimics this expression internally. Egermann [8] also uses the term *Contagion* to describe an unconscious automatic mimicking of others' expressions affecting ones' own emotional state. We take this theory further and hypothesize that, as it is possible for preceding music to contagiously affect not only the listener's emotion, but also the perception of emotions, thereby resulting in modified perception of the emotions of the musical pieces or phrases that follow. This is drawn from the basic fact that emotions are contextual – thus the meaning making of one music clip might depend on the musical context it is heard in. This approach may explain why musical emotions are perceived in the way they are perceived – the same musical notes make us happy in one song and sad in another, depending on the notational and lyrical context of the musical piece.

1.2 Capturing Human Emotions

The most common method of capturing emotions is self reporting. In the context of musical emotions, the works of Egermann and Schubert are pioneering. Egermann [8] presented the concept of web based surveys, where data regarding emotion, empathy, and preference ratings were gathered from a music-personality test, which had twenty three musical excerpts as stimuli. Schubert et. al [9], [10] created an interface based on facial expressions (emoticons) "aligned in clock-like distribution", through which survey participants could quickly and easily rate emotions in music continuously. The emoticons they used depicted six emotions: Excited, Happy, Calm, Sad, Scared and Angry. We drew from these works to construct an interface that makes it possible to record people's emotions as they listen to music discretely and continuously – this is discussed in details in section 2.2. Continuous emotion rating of music has resulted in better understanding and modeling of music's emotion-generating nature. This also helps in mapping musical features, structures and constructs onto emotional response data [11], [12].

1.3 Emotions in HCM - Nāva Rāsa

Our study focuses mainly on perception of *Hindustani Classical Music*, abbreviated as *HCM*. *HCM* is one of the two major branches of *Indian Classical Music* (*ICM*), the other being *Carnatic Music*. The basic melodic structure of HCM is the *Rāga*, which, according to Chordia [13] and others "is a collection of melodic atoms and a technique for developing them. These melodic atoms are sequences of notes that are inflected with various micro pitch alterations and articulated with expressive sense of timing. Longer musical phrases are built by knitting these melodic atoms together". The tempo in HCM are primarily of three types - *Vilāmbit* (slow), *Mādhyalāya* (a little faster), and *Drut* (fast). It might be noted that one significant aspect of HCM is the practitioner's freedom of improvisation, keeping the grammar of Rāga and the tempo in mind. Many of the *Rāga*s are said to be associated with particular emotions, which are elicited through the use of different melodic constructs (or features). In Indian aesthetics, emotional responses to any art form are said to be one of nine prevalent types, or *Nāva Rāsa*, which literally translates to Nine (= *Nāva*) Emotions (= *Rāsa*). These are *Romantic*, *Happy*, *Wonder*, *Calm*, *Anger*, *Courage*, *Sad*, *Fear* and *Disgust*. These are the emotions that a human perceives/shows in response to different situations, according to Bhārāta Muni's treatise *Nātyashāsthra* [14]. The concepts of *Nāva Rāsa* have influenced our survey interface, which is described in section 2.2.

In this paper, we study the phenomenon of *perceptual contagion of perceived emotion*, in the context of music presented in sequence, using a novel contagion capturing interface – which is based on the continuous music survey framework, with the aim to understand if and how the perception of one music clip biases the perception of the next music clip. To the best of our knowledge, this study is the first of its kind, with HCM as a special case study.

2 Surveys and Data Acquisition

In this section we describe the survey structure and interface we used to capture people's emotions which were evoked when they were exposed to excerpts or clips of various music. Around 300 students belonging to different courses of IIT Kharagpur, India, took part in the surveys. The age range varied from 18 to 30 years. As part of the survey, demographic details like gender, music preference, musical training were also recorded. It might be noted here, that this is possibly the first time that such a large scale survey for understanding HCM's emotive aspects have been undertaken.

2.1 Stimuli

We used three sets of clips for the present work. The first set – called **IM1_Clips** – constituted of the six international clips used by Schubert et al. [10]. Since these clips were already annotated with their emotional content in the literature [10], we validated the correctness of our survey by comparing our ratings with Schubert's. The context-dependent biases for perception of some of these clips were also verified. The prefix *IM* in the clips' names stand for *International Music*. The number following *IM* indicates set number. The second and third sets constitute of HCM clips, called **HCM1_Clips** and **HCM2_Clips**, which were chosen by us. Each HCM clip was first annotated with the major emotion(s) it elicits by experts. These emotions depend on a number of factors like the major and minor note sequences of the clips and tempo. Generally, the use of major and minor notes elicit positive and negative valences respectively. Universally, high and low tempos are indication of high and low arousal respectively. The details of all the clips are provided in table 1. It might be noted that perception of emotion is not based on familiarity of the music clip, but on intrinsic musical features.

Table 1. Details of Clip–sets used in our surveys

Clip_Name	**Origin**	**Clip_Emotion**	**Clip_Duration (Seconds)**
IM1_Clips			
IM1_Exc	Toy Story: Infinity and Beyond	Exciting	16
IM1_Hap	Cars: McQueen and Sally	Happy	16
IM1_Calm	Finding Nemo: Wow	Calm	16
IM1_Sad	Toy Story 3: You Got Lucky	Sad	21
IM1_Fear	Cars: McQueen's Lost	Fear	11
IM1_Angry	Up: 52 Chachki Pickup	Angry	17
HCM1_Clips			
HCM1_Hap	Hamsadhwani	Happy	40
HCM1_Calm	Komal Rishabh Asavari	Calm	45
HCM1_Sad	Marwa	Sad	47
HCM2_Clips			
HCM2_Hap1	Desh	Happy	46
HCM2_Hap2	Desh	Happy	52
HCM2_Calm	Bhairavi	Calm	48
HCM2_Sad1	Marwa	Sad	40
HCM2_Sad2	Marwa	Sad	50

2.2 Survey Description

The objective of our surveys was to find out – through the method of self report – how and if, the cognitive perception of one music clip gets biased due to the perception of the clip preceding it, when listeners are subjected to music clips conveying similar or radically different emotions. It might be so that a music clip elicits more than one type of emotion in the listeners. So, the interface was so designed that users may choose multiple emotions over the duration of a single clip. The survey interface consisted of a wheel structure, as seen in figure 1. The wheel has three distinct parts: *Central part* – which contains the play and pause buttons for the music clips. *Neutral part* – The cursor is supposed to rest here while a person decides on the emotional response. The *10 spokes* – Eight spokes represent eight emotions inspired by the *Nāva Rāsa* concept, a) Happy (H), b) Calm (C), c) Wonder (W), d) Exciting (E), e) Anger (A), f) Fear (F), g) Sad (S), h) Romantic (R). *Other Emotion* (OE) represents an emotion which is absent in the wheel and *Don't Know* (DK) represents the user's indecisiveness regarding the emotional response. Each spoke is divided into 5 sub-regions, which indicate the intensity levels for each emotion. The inner-most region, which has the lightest shade of color, marks the lowest intensity. Both shade and intensity increase as we move outwards. The survey taker is expected to click on their perceived emotion intensities, any number of times. The website used HTML, PHP scripts and MySQL in the back-end. The survey results stored in the database were then processed using MATLAB.

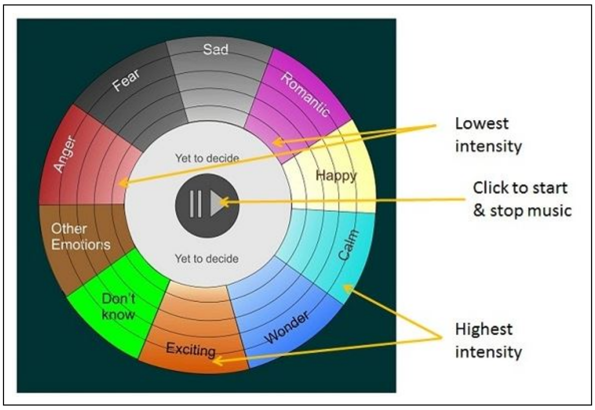


Fig. 1. The Web–Based Survey Interface

2.3 Types of Surveys

In this study two major types of surveys were used, a) Discrete Music Rating (DMR) and b) Contagion Interface Rating (CIR).

Discrete Music Rating Surveys (DMR) : In all the surveys, survey takers are subjected to more than one music clip. The idea in DMR is to present each music clip discretely, that is, with a perceptible time gap between clips. There may be questions (demographic, and/or related to the music clips) in between two music clips. This helps in removing any cognitive biases that might have been formed in the survey taker, and helps us to get the unbiased emotion ratings for each clip. The sequence of the clips were random. It might be argued that the true emotional contents of the music clips are obtained through DMR surveys. These findings are compared to the biased ratings for the same clips, to understand the extent and patterns of biases.

Contagion Interface Rating Surveys (CIR) : Users taking CIR surveys listen to a series of different music clips without any time gap between two consecutive clips. The music clips in the survey are of a pre-defined emotion sequence: From positive emotion generating clips to negative emotion generating clips, and vice versa. For example: The sequence of clips can be like HappyClip $\rightarrow$ CalmClip $\rightarrow$ SadClip and its opposite sequence.

2.4 Methodology and Data Normalization

For each clip, the data coming from survey has a rich temporal structure of user clicks on various emotions with different intensities. In this work, we are interested in studying effects of one clip on the perception of next. For this, we compute a uniform representation of various emotions for each $< user - clip >$ pair. We call this the *normalized representation*, which loses the temporal pattern of clicks, but is easily interpretable for the current study. For a clip c, a user u can click on any emotion - any number of times. Assuming he has clicked $n_x^{c,u}$ times on the emotion (x), where $x \in \{DK, OE, F, A, S, C, W, R, H, E\}$, the complete list of emotions used in survey (section 2.2). Thus, the total number $N^{c,u}$ of clicks by user u for clip c is:

$$N^{c,u} = n_{DK}^{c,u} + n_{OE}^{c,u} + n_F^{c,u} + n_A^{c,u} + n_S^{c,u} + n_C^{c,u} + n_W^{c,u} + n_R^{c,u} + n_H^{c,u} + n_E^{c,u} \quad (1)$$

Let each of these $n_x^{c,u}$ responses have intensities $I_x^{c,u}(1)$, $I_x^{c,u}(2)$, ..., $I_x^{c,u}(n_x^{c,u})$, where $1 \leq I_x^{c,u}(k) \leq 5$ and $1 \leq k \leq n_x^{c,u}$. We calculate average intensity for each emotion as $\bar{I}_x^{c,u}$, as:

$$\bar{I}_x^{c,u} = \frac{\sum_{k=1}^{n_x^{c,u}} I_x^{c,u}(k)}{n_x^{c,u}}. \quad (2)$$

The normalization procedure should allow us to take into consideration two measures. First, the number of clicks on a particular emotion, with respect to total clicks for that clip should be considered, so that each user gets one normalized vote regarding an emotion in the clip, despite using the freedom to click that emotion as many times as he wants. This measure normalizes the votes of an over-active survey taker (who clicks many times) with that of one who clicks moderately. This will give an indication of frequency of the emotion felt by the user. So, greater values of $\left(\frac{n_x^{c,u}}{N^{c,u}}\right)$ – where $0 \leq \left(\frac{n_x^{c,u}}{N^{c,u}}\right) \leq 1$ – indicate that the clip c elicited the emotion x in the user u, more than any other emotion. Second, the intensity average of a particular emotion should be considered. This provides a measure of how intensified the feeling of a particular emotion was. A high $\bar{I}_x^{c,u}$ value will indicate greater feelings of emotion x over the clip c, by user u. Therefore, we define the weight of emotion x, in clip c, for user u as:

$$Wt_x^{c,u} = \left(\frac{n_x^{c,u}}{N^{c,u}}\right) * \bar{I}_x^{c,u} = \frac{\sum_{k=1}^{n_x^{c,u}} \bar{I}_x^{c,u}}{N^{c,u}} \quad (3)$$

The weight of each emotion x, for each clip c, rated by the user u can be calculated from equation 3. Considering entire survey results, for each clip, we aggregate the weights for each emotion over all users to derive the ultimate percentage emotion contents of that clip.

3 Results and Discussion

In this section, the results of the various surveys taken are discussed. Our work is divided into two parts – a) Clustering the clips according to their dominant emotions and b) Observing how the perception of the clips of different clusters are biased across the surveys.

Table 2 gives the details of the DMR surveys, from which the dominant emotion of each clip can be identified. The *Dominant Emotion* of a clip is the emotion which has highest percentage value from the ratings across the surveys. For each clip, these are marked bold in the table. These results show that the interface is able to capture the emotional contents of the clips well. The results of DMR-1 correspond to Schubert's [10] findings of the dominant emotions for the clips. The results of DMR-2 and DMR-3 comply with the expected dominant emotions of the clips.

From the DMR surveys, we group the clips into three clusters according to their dominant emotions – a) **Happy clips** – IM1_Hap, HCM1_Hap, HCM2_Hap1, HCM2_Hap2, b) **Sad Clips** – IM1_Sad, HCM1_Sad, HCM2_Sad1, HCM2_Sad2 and c) **Calm Clips** – IM1_Calm, HCM1_Calm, HCM2_Calm. These clips are further used in the CIR surveys, to observe if and how the perception of clips of each cluster are affected by the context it is heard in. As explained earlier, the sequence of clips in the CIR surveys are predefined and fixed, which are given in table 3.

Table 2. Setwise DMR results for all the clips

Clip_Name	DK%	OE%	F%	A%	S%	C%	W%	R%	H%	E%	Dominant Emotion
DMR-1											
IM1_Exc	0	0	23.03	11.60	6.46	9.08	0	0	20.92	**28.90**	**Exciting**
IM1_Hap	0	0	5.33	6.29	10.59	17.45	0	5.62	**29.69**	25.03	**Happy**
IM1_Calm	0	0	3.67	3.04	19.53	**39.94**	11.79	10.93	8.54	2.56	**Calm**
IM1_Sad	0	0	3.39	3.08	**29.88**	25.52	18.53	12.36	3.36	0	**Sad**
IM1_Fear	0	0	**30.74**	24.35	9.59	0	0	0	10.69	24.63	**Fear**
IM1_Angry	0	0	24.37	**31.39**	6.37	3.88	0	0	6.54	27.45	**Angry**
DMR-2											
HCM1_Hap	0	0	1.58	1.98	4.95	21.15	15.61	10.87	**24.3**	19.56	**Happy**
HCM1_Calm	0	0	5.79	2.18	23.41	**29.43**	14.09	12.94	6.52	6.04	**Calm**
HCM1_Sad	0	0	10.95	3.24	**38.1**	28.62	8.55	5.57	1.99	2.98	**Sad**
DMR-3											
HCM2_Hap1	0	0	2.02	2.23	7.59	22.99	14.51	12.73	**26.76**	18.67	**Happy**
HCM2_Hap2	0	0	1.05	2.06	2.49	15.98	15.15	17.84	**23.43**	14.5	**Happy**
HCM2_Calm	0	0	5.51	1.94	25.71	**34.15**	8.81	8.55	7.71	3.3	**Calm**
HCM2_Sad1	0	0	8.71	4.75	**30.03**	20.57	10.84	11.87	12.13	9.5	**Sad**
HCM2_Sad2	0	0	12.72	8.4	**33.33**	20.85	10.91	6.3	3.83	3.83	**Sad**

Table 3. Clip sequences for the CIR surveys

Survey Name	Clip Sequence
CIR-1_PosNeg	IM1_Hap → IM1_Calm → IM1_Sad
CIR-1_NegPos	IM1_Sad → IM1_Calm → IM1_Hap
CIR-2_PosNeg	HCM1_Hap → HCM1_Calm → HCM1_Sad
CIR-2_NegPos	HCM1_Sad → HCM1_Calm → HCM1_Hap
CIR-3_PosNeg	HCM2_Hap1 → HCM2_Hap2 → HCM2_Calm → HCM2_Sad1 → HCM2_Sad2
CIR-3_NegPos	HCM2_Sad2 → HCM2_Sad1 → HCM2_Calm → HCM2_Hap2 → HCM2_Hap1

Table 4 gives the results of all the CIR surveys. The clusters of clips with same dominant emotion are shown together so that their biases may be compared.

Table 4. Setwise CIR survey results for the clips

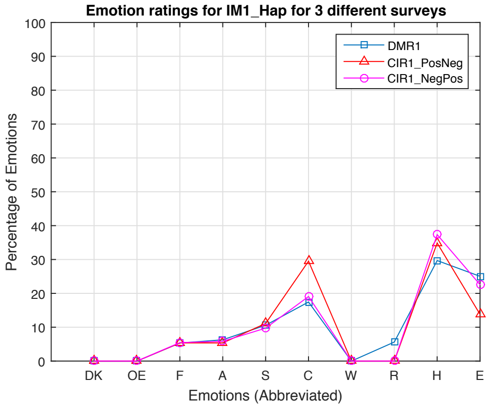
Clip_Name	Survey_Type	DK%	OE%	F%	A%	S%	C%	W%	R%	H%	E%
Happy Clips											
IM1_Hap	PosNeg	0	0	5.36	5.36	11.11	29.51	0	0	**34.87**	13.79
	NegPos	0	0	5.45	5.84	9.73	19.07	0	0	**37.35**	22.57
HCM1_Hap	PosNeg	0	0	1.76	2.03	4.07	22.10	4.45	9.32	**31.89**	24.38
	NegPos	0	0	0	0	1.59	4.52	4.91	10.17	**48.00**	30.81
HCM2_Hap1	PosNeg	0	0	0	0.68	0	19.37	4.33	2.28	**39.60**	33.70
	NegPos	0	0	1.40	2.72	0.49	2.12	3.28	7.22	**50.07**	32.63
HCM2_Hap2	PosNeg	0	0	1.36	4.23	2.27	19.06	7.23	8.95	**27.70**	29.20
	NegPos	0	0	0.09	0.68	0.09	1.51	3.57	14.85	**61.36**	17.86
Calm Clips											
IM1_Calm	PosNeg	0	0	6.45	6.45	21.37	**36.58**	0	0	14.52	9.27
	NegPos	0	0	5.36	5.36	11.11	29.51	0	0	**34.87**	13.79
HCM1_Calm	PosNeg	0	0	1.25	0	**67.08**	19.29	2.36	8.38	1.64	0
	NegPos	0	0	1.24	0	25.69	**29.55**	19.33	6.22	13.82	4.15
HCM2_Calm	PosNeg	0	0	2.07	2.9	**58.68**	14.50	15.87	0	6.03	0
	NegPos	0	0	7.23	3.41	21.63	**35.96**	11.45	4.01	11.45	0.69
Sad Clips											
IM1_Sad	PosNeg	0	0	9.87	8.97	**47.53**	17.94	0	0	7.62	8.07
	NegPos	0	0	9.13	6.22	**33.76**	31.12	0	0	14.75	8.71
HCM1_Sad	PosNeg	0	0	10.79	0	**56.63**	21.45	5.84	3.29	0	0
	NegPos	0	0	0	0	**47.73**	30.04	14.86	0.56	3.40	3.40
HCM2_Sad1	PosNeg	0	0	23.30	5.18	**44.87**	11.50	12.96	7.44	0	0.68
	NegPos	0	0	5.05	3.20	**38.90**	23.10	3.73	1.26	12.69	6.10
HCM2_Sad2	PosNeg	0	0	20.60	1.42	**42.31**	12.3	11.12	7.26	2.65	2.36
	NegPos	0	0	3.43	0.12	**45.47**	43.54	4.44	2.22	0.12	0.66

3.1 Comparison of DMR and CIR Values

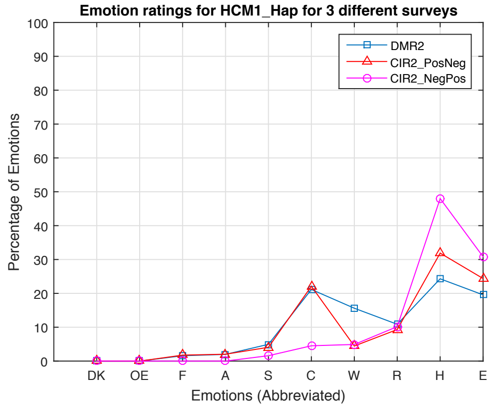
From the above surveys, we were able to ascertain the dominant emotion contents of each clip, from the maximum percentage values of the emotion vectors. Depending on these contents, we grouped the clips into 3 distinct categories – a) *Happy Clips* – which express positive emotions, like happiness, excitement. b) *Sad Clips* – which express negative emotions, like sadness. c) *Calm Clips* – which express calmness. For each clip, 3 surveys provide data on how the perception changes under different context. This is discussed below for each group of clips.

Happy clips The dominant emotion for the clips of this set is Happy (H). The major observations for this set are drawn from the comparative graphs in figure 2, which depicts the perceptual contagion of emotion for all Happy clips, over all surveys, when they are preceded by Happy, Sad and Calm clips.

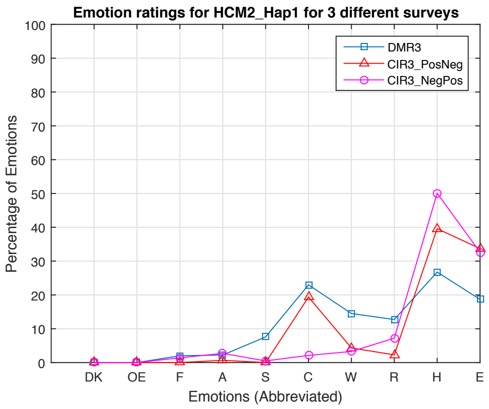
1. Depending on the survey type, significant differences in values for emotions *Happy*, *Exciting* and *Calm* were observed. The values of H and E are least in the base ratings, captured in the DMR surveys. In the CIR_PosNeg surveys, H and E increase slightly than the base values. In the CIR_NegPos surveys, H and E increase significantly than the base values. We may conclude that when happy music follows sad and calm music, it might be perceived as even more happy.
2. All happy clips have a *Calm* component. The Calm values are highest in CIR_PosNeg surveys. They go down significantly in the CIR_NegPos surveys, where the Positive clips are perceived as more positive.



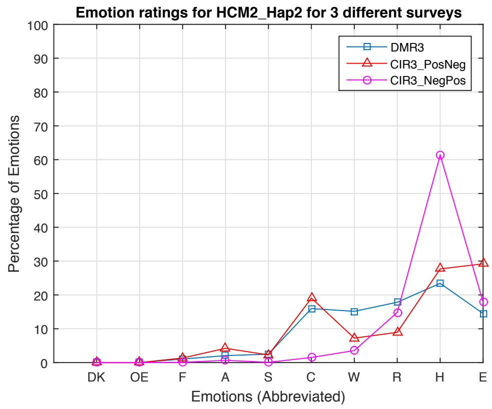
IM1_Hap



HCM1_Hap



HCM2_Hap1

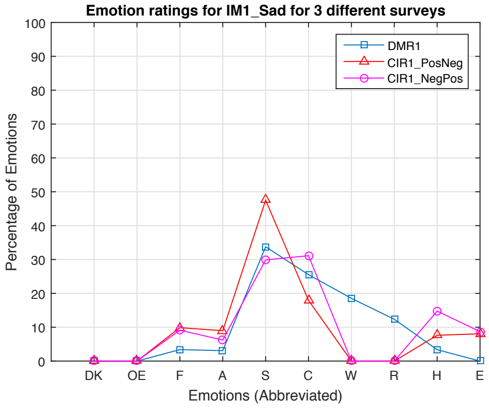


HCM2_Hap2

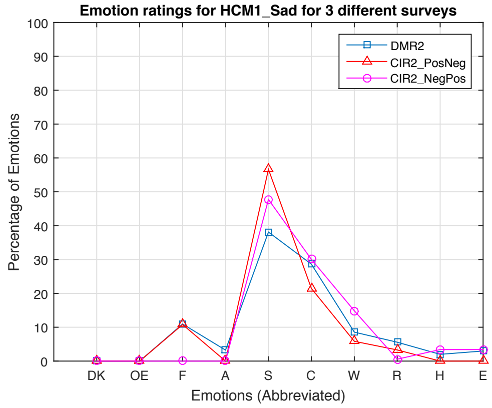
Fig. 2. Perceptual Contagion of Emotion for all Happy clips.

Sad Clips The dominant emotion for the clips of this set is Sad (S). The major observations for this set are drawn from the comparative graphs in figure 3, which depicts the perceptual contagion of emotion for all Sad clips, over all surveys, when they are preceded by Sad, Happy and Calm clips.

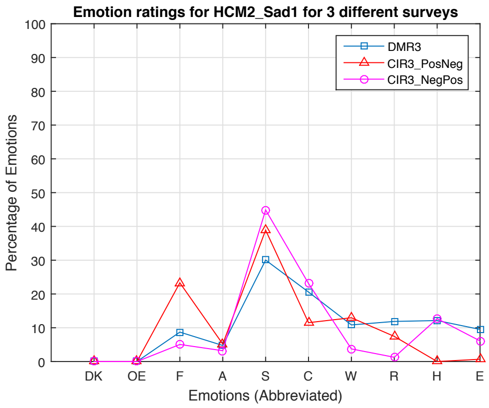
1. Depending on the survey type, significant differences in values for emotion *Sad* was observed. The emotion values of S are least in the base ratings, captured in the DMR surveys. In the CIR_PosNeg surveys, S values increase significantly more than the base values. In the CIR_NegPos surveys, S values are nearer to the base values. We may conclude that when sad music follows happy and calm music, it might be perceived as even more sad.
2. The Calm (C) values are very low in CIR_PosNeg surveys. They are very high in the CIR_NegPos surveys, where the sad clips are perceived as almost baseline sad.



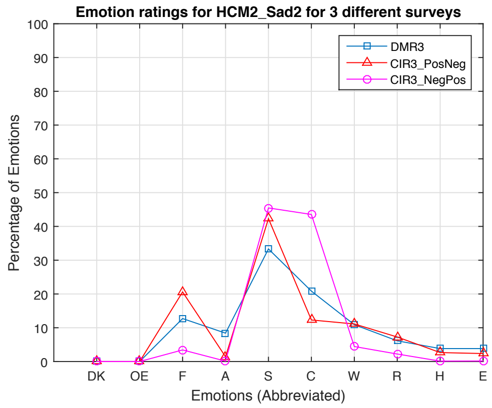
IM1_Sad



HCM1_Sad



HCM2_Sad1



HCM2_Sad2

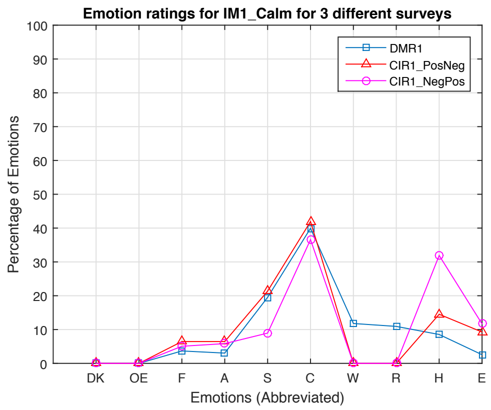
Fig. 3. Perceptual Contagion of Emotion for all Sad clips

3. Fear (F) contributes in the CIR_PosNeg surveys significantly for all the clips in this set.

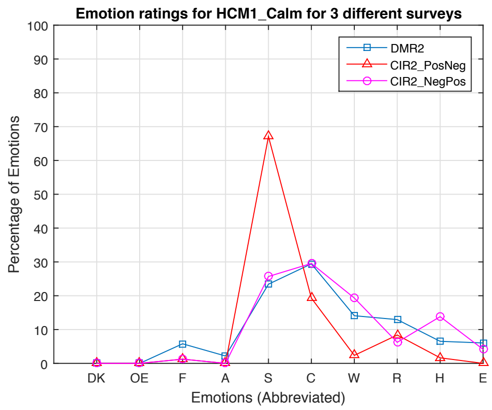
Calm Clips The dominant emotion for the clips of this set is Calm (C). The major observations for this set are drawn from the comparative graphs in figure 4, which depicts the perceptual contagion of emotion for all Calm clips, over all surveys, when they are preceded by Sad and Happy clips.

1. Depending on the survey type, significant differences in values for the emotion *Calm* was observed. The baseline values of C are captured in the DMR surveys. C values go significantly up in the CIR_NegPos surveys. In the CIR_PosNeg surveys, C decreases drastically than the baseline values. We may conclude that when calm music follows happy music, it might be perceived less calm and when it follows sad music, it might be perceived as more calm.

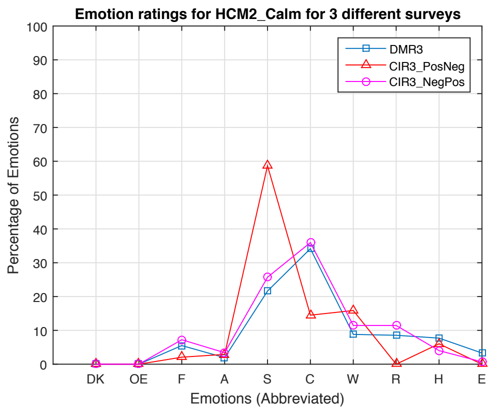
2. Interestingly, the calm clips are perceived as very sad in the CIR_PosNeg surveys. Sometimes, the S component was even higher than the C component in these surveys. Again, in the CIR_NegPos surveys, the S component were significantly low.
3. High Happy (H) component was observed for these clips in the CIR_NegPos surveys, where the S component was significantly low.



IM1_Calm



HCM1_Calm



HCM2_Calm

Fig. 4. Perceptual Contagion of Emotion for all Calm clips

4 Conclusion

Results suggest that, perceived emotions of musical clips are dependent on the emotional content of the preceding clips. Also, the type of dependence and how the perception changes depend on the emotional content of the musical clip as well. Happy clips may be perceived as far more happy than they actually are, when they are heard in a sequence, after some sad and calm clips. Again, sad clips may be perceived as far more sad than they actually are, when they are heard in a sequence, after some happy and calm clips. Calm music is perceived to be much more calm and happy, when heard after sad clips, but are perceived to be much less calm and even sad, when heard after happy clips. These trends suggest that over short durations, perception of musical emotions are modified by what we have heard just earlier. Hence, music listening is perceptually modified by earlier musical phrases and their emotional contents. Future research may need to address the exact nuances of music to music perceptual contagion, their possible reasons, and implications for music appreciation.

References

1. Berkowitz, L.: On the formation and regulation of anger and aggresion. A cognitive-neoassociationistic analysis. The American Psychologist. 45(4), 494-503, (1990).
2. Frijda, N. H.: The Emotions. Cambridge University Press, (1986).
3. Keltner, D., Buswell, B.N., : Embarrassment: its distinct form and appeasement functions. Psychological bulletin, 122(3), (1997).
4. Ekman, P. : An argument for basic emotions. Cognition & emotion, 6(3-4), 169-200, (1992).
5. Juslin, P. N., Sloboda, J. A. : Music and emotion. The psychology of music, 3, 583-645, (2001).
6. Juslin, P. N.: From mimesis to catharsis: expression, perception, and induction of emotion in music. Musical communication, 85-115, (2005).
7. Juslin, P. N.: From everyday emotions to aesthetic emotions: towards a unified theory of musical emotions. Physics of life reviews, 10(3), 235-266, (2013).
8. Egermann, H., McAdams, S.: Empathy and emotional contagion as a link between recognized and felt emotions in music listening. Music Perception: An Interdisciplinary Journal, 31(2), 139-156, (2013).
9. Schubert, E., Ferguson, S., Farrar, N., Taylor, D., Mcpherson, G. E.: Continuous response to music using discrete emotion faces. Proc. of CMMR, (2012).
10. Schubert, E., Ferguson, S., Farrar, N., Taylor, D., McPherson, G. E.: The six emotion-face clock as a tool for continuously rating discrete emotional responses to music. In International Symposium on Computer Music Modeling and Retrieval (1-18). Springer Berlin Heidelberg, (2012, June).
11. Schmidt, E. M., Turnbull, D., Kim, Y. E.: Feature selection for content-based, time-varying musical emotion regression. In Proceedings of the international conference on Multimedia information retrieval (267-274). ACM, (2010, March).
12. Schubert, E.: Continuous self-report methods. (2011).
13. Chordia, P., Rae, A.: Raag Recognition Using Pitch-Class and Pitch-Class Dyad Distributions. In ISMIR (431-436), (2007, September).
14. Bharatmuni: Natyasastra of Bharata, Chapter 6.