2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Affective response to tunes synthesized with musical pitch curves

Venkata S Viraraghavan, Tince Varghese, Rahul Dasharath Gavas, Mithun B. S,
Ramesh K Ramakrishnan, Balamuralidhar P, Arpan Pal
TCS Research, India

***Abstract*— Tunes perceived as happy may help a user reach an affective state of positive valence. However, a user with negative valence may not be ready to listen to such a tune immediately. In this paper, we consider nudging a user from their current affective state to a target affective state in small steps. We propose a technique to generate a gradation of tunes between an initial-reference tune and a target-reference tune, to achieve the affect transition. The two-dimensional gradation is realized in time and in pitch, respectively, by varying the tempo and by the use of musical pitch curves, i.e. pitch transients or simply 'transients'. We exploit the duration and scaling of transients observed in South Indian music (Carnatic) to introduce transients into existing tunes. In our experiment, we have introduced the transients into Western music tunes. The results of perceptual evaluation show that the affective response to transients is likely to be higher at slow tempos than at fast tempos. Further, when felt, transient-tunes are twice as likely to be associated with positive valence than with negative valence, irrespective of tempo.**

I. Introduction

Emotional well-being, an important aspect of mental well-being, is the emotional aspect of everyday experience [1]. Recently, several organizations and enterprises have increased attention on emotional well-being at the workplace (e.g. [2]). The impact of music on the human mind has been studied for several centuries and has applications in therapy [3]. An important application in the enterprise-context is the use of music as an intervention for changing emotional states (e.g. Destressing). Consider a closed-loop intervention system such as in Fig. 1. In this type of system, the user's mental state is sensed by physiological signals. The typical signals are voice, image, video, photoplethysmogram, temperature, etc. in a mobile/wearable based system, while electroencephalogram (EEG), etc. may also be used in laboratory conditions. The mental state is often assessed using Russel's two-dimensional model of affect, where affect is constituted by two independent axes called valence and arousal. The affect Happy is an example of positive valence and Sad is an example of negative valence. Similarly, Angry is an example of high arousal (negative valence) and Contented is an example of low arousal (positive valence). In the system of Fig. 1, if the sensed valence remains negative for, say, two days, an intervention can be triggered. One possible intervention is to play appropriate music to nudge the user towards positive valence. The music used in intervention are usually well-known creations and often subject-specific. In Europe, Western classical music was an oft-considered genre. Other modern genres including user-preferred genres may also be used [4]. Therapy involving Indian art music (IAM) is also gaining ground. While there is choice among the genres of music, the choice in the components of music has largely been ignored. Just as components of Western music (WM) have been studied (e.g. chords [5]), the components of Indian music also need to be studied [6]. Common elements between WM and IAM do exist,

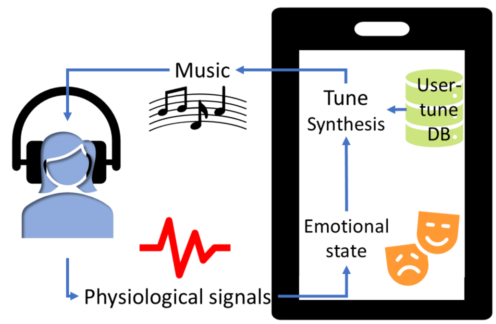

Fig. 1. Intervention system architecture

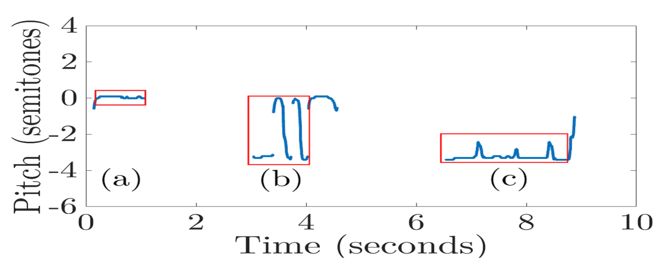

Fig. 2. Examples of (a) constant pitch segment and (b,c) deliberate transients in IAM

but several components of IAM are different from those of Western music. We explain the difference using the example of a piano and violin. Piano keys are tuned to discrete pitches (or fundamental frequencies) of musical notes and pitches 'in between' cannot be produced. In a violin, however, pitches between those of musical notes can be produced by changing the position of the finger. Further, the pitch can be varied continuously by sliding the finger even as the bowing is sustained. We call the corresponding musical pitch curves as pitch-transients, or henceforth, 'transients'. Transients such as glissando and portamento are optional in Western music. In IAM, transients are mandatory and carry information in vocal and instrumental music [7]. In Fig. 2, an instance of a fixed pitch is marked as (a), while two instances of transients in IAM are marked as (b) and (c). Such transients are taught and learnt carefully. By contrast, harmony is used profusely in WM, but has a negligible role in IAM. In this paper, we propose the use of transients in music for intervention based on Carnatic Music (CM) [8], which is a well-known sub-genre of IAM. However, the proposal is not limited to CM.

A. Related work

It has been found that tunes in the major mode (Ionian mode) are perceived as happiest in comparison to other modes [9]. There is also evidence to suggest that the emotional responses to music are dominated by the melody rather than by lyrics [10]. In fact, it appears that adult listeners can interpret the emotion conveyed in music of familiar and unfamiliar cultures according to their perception of acoustic cues [11]. These studies follow the 'cognitivist' view of musical emotion, where music is expected to convey an emotion. The 'emotivist' view is that music makes us feel the emotion [9]. Emotivist studies indicate that sad music can indicate pleasant emotion [12] and thus, there is a need for subjective solutions [13]. Music therapy and intervention systems (such as in Fig. 1) need to consider both views. A pre-requisite is to understand the emotional impact of genres of music being employed and their components. For example, in WM, emotion-related brain areas are activated even by a single musical chord [5]. The impact of IAM is usually studied in relation to *rāga*s. The authors of [14] used fMRI to used fMRI to measure the response, while [15] monitored EEG signals of listeners. Since some results question the *rāga*-based approach [15], we pursue a component-based approach. Although the components of IAM have been researched, their emotional impact needs to be understood. Specifically, they have not been used in experiments that involve physiological sensing (such as the EEG, fMRI etc.) to understand their impact. The deliberate use of transients is thus not viable yet for the music therapist or in an intervention system. Since transient-related research on emotion is nascent, it is necessary to first obtain subjective feedback. Such feedback can guide further research that may use physiological signals to verify perceptions. In this paper, we study the emotional impact of transients added automatically to the main melody track of Western tunes.

II. Methodology

The intervention system mentioned in Section I serves to motivate our proposed methodology. Consider a user whose mental state is Sad. A possible intervention is to play a tune that can change the mental state of the user to Happy. We also expect that a user who is currently Sad may not engage with a tune that is clearly perceived as Happy. A possible solution is to start with tunes that are perceived as Neutral and change to tunes to those that are perceived as more Happy, and finish with a *target-reference tune* that is clearly perceived as Happy. However, this involves changing tunes, which may also detract the user from continuing to listen to the intervention. Further, a real system will need as many subjective evaluation results as there are tunes. We therefore seek a technique where the user feels that each tune has changed very little compared to the previous one. As the tunes change, it reaches the target-reference tune. If each of these progressive tunes is generated from the target-reference tune, we need subjective information about only this tune and the impact of the generation mechanism. Currently, a well-known approach is to change the speed of the tune (which we call 'gradation in time'). We propose below a two-dimensional technique that also includes a novel 'gradation in pitch' through the use of transients.

A. Generation of intervention tunes

We restrict our description to the typical case where major-scale notes are perceived as Happy and minor-scale notes are perceived as Sad [16]. Since such associations are subjective, a personalized intervention system is necessary in practice. Further, the gradation technique below can be extended to any pair of tunes.

Consider a monophonic tune that is known to be perceived as Happy. Examples of two such tunes are the main melody tracks of the two Western compositions, which are known to be associated with happiness [16]. Since this tune is associated with Happy, we treat this as the target-reference tune of the audio intervention. We further hypothesize that the same tune at a much slower tempo, and with the major-scale notes replaced by minor-scale notes, is closer to Sad than Happy. We call this tune the *initial-reference tune*, which may be generated in this manner, or by any other means which may take subjective feedback into account. Starting from the initial-reference tune (typically perceived as Sad), we introduce gradation along the two axes of time and pitch to generate a series of successive tunes that leads to the target-reference tune (typically perceived as Happy).

B. Gradation in time

Let the tempo (i.e. speed) of a given tune, $\mathcal{T}$, be R beats per minute. We generate N intermediate tunes, which are increasing in tempo. In this paper, we use a linear, one-dimensional gradation in time, which is realized by varying the tempo in N linear steps from $\frac{R}{2}$ to R. The time-gradation factor τ of a tune is the ratio of its tempo to R. For the initial-reference tune, we set $\mathcal{T}_1$, $\tau = 0.5$ and for the target-reference tune, we set $\mathcal{T}$, $\tau = 1.0$. Thus, the duration of each note/silence-segment in the initial-reference tune is twice the duration of the corresponding note/silence-segment in the target-reference tune. For the intermediate tunes, the duration scale inversely as τ.

C. Proposed gradation in pitch

We use transients to introduce a gradation in pitch between the initial- and target-reference tunes. Consider the 12 notes in an octave starting from a reference note, called the key in WM. Note 0, which is the key itself, and Notes 1 to 11 constitute the octave. The major-scale consists of Notes 0, 2, 4, 5, 7, 9, 11. The minor-scale consists of Notes 0, 2, 3, 5, 7, 8, 10. To obtain the initial-reference tune, we first find all occurrences of Notes 4, 9 and 11 with respect to the key specified for the target-reference tune. We then replace each of these notes with the corresponding minor-scale notes, Notes 3, 8 and 10, respectively. Next, let the duration of an original note be L_N. Then, the duration of the replaced note is also L_N. To introduce pitch gradation, each replaced note is converted to a pitch curve of duration L_N. This pitch curve consists of an anchor note (of fixed pitch) and a pitch transient that follows the anchor.

While several choices of anchor and transient-duration are possible, we borrow from observations on CM [17] in this paper. Accordingly, Notes 2, 7 and 9 serve as the anchors for Notes 3, 8 and 10, and the nominal duration of a transient is $L_T = 300$ms. We introduce transients from the anchor-pitch, a, to q notes above a. We use large transients by setting $q = 3$ (instance (b) in Fig. 2), to cater to listeners that are not familiar with CM. However, there are other possible choices: e.g. the transient may be one note lower than the anchor, where $q = -1$. The duration of the transient is fixed as follows. If $L_N \leq L_T$, the pitch curve $p(t)$ consists of only a transient of duration L_N. If $L_N > L_T$, an anchor of duration $l_a = L_N - L_T$ precedes the transient. Thus:

$$p(t) = a, 0 \leq t < l_a \quad (1)$$
$$= a + q\cos(\pi \frac{t - t_0}{l_T}), l_a \leq t < l \quad (2)$$

In equation 2, cosine-curves are used to model the transients. Other models are possible and are generally indistinguishable even by CM experts [18]. Just as τ quantifies the time-gradation, we use ρ, to quantify gradation in pitch. The quantity ρ is defined as the duration of major-scale notes in the pitch curve, relative to the duration of the replaced notes. For the initial-reference tune, $\rho = 0$, and the for target-reference tune, $\rho = 1.0$. For tunes with transients, ρ depends on the tempo. For a given tempo, the target-reference tune, the generated initial-reference tune, and the generated transient tune are shown in Fig. 3.

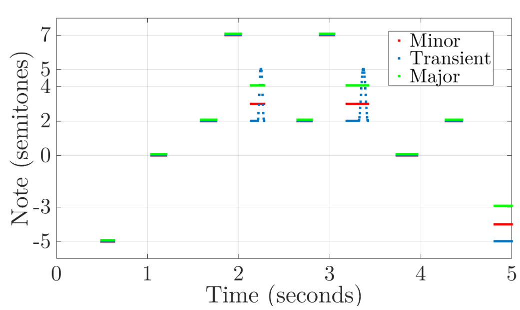

Fig. 3. Note-vs-time curve (Transient, blue) for a tune. The initial-reference tune (red) and the target-reference tune (green) are also shown.

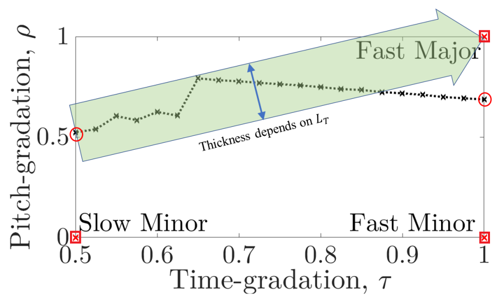

Fig. 4. Realization of pitch gradation, i.e. ρ, for $N = 20$ gradations in time for the two tunes given in Section II.

D. Two-dimensional gradation

Time and pitch gradation combine as shown in Fig. 4. Tunes without transients are marked by a crossed check box (☒) and the tunes with transients are marked by circles. For each time gradation, the pitch gradation value that can be realised is plotted as a dotted line. Exactly one pair of time and pitch gradation is possible due to the constraint imposed by $L_T = 300$ ms. If this constraint is relaxed fully, the entire region in the pitch/time gradation area can be realised. However, not all tunes generated thus are perceived as musical [19]. In practice, for musical tunes, a smaller region should be covered. An example is shown by the thick arrow whose width is governed by the range of L_T. This arrow also shows an intervention path made possible by our proposed gradation. Without pitch gradation, the only paths possible are horizontal lines at $\rho = 0$ and $\rho = 1$.

III. Evaluation and Results

We used the pitch-bend commands through a Musical Interface for Digital Instruments (MIDI) [20] to synthesize the pitch curves, generated according to Section II. To understand the effect of gradation, we conducted a listening experiment. In such experiments, participants tend to remember the tunes they heard earlier [21]. Thus, we chose only specific possibilities from nearly 50 tunes.

Two happy tunes are chosen for the experiment based on the results in [16]. For each tune, its monophonic melody track is referred to as the 'Fast-major' flavor. This flavor is at the original tempo, but is called 'Fast' because happy tunes are typically fast. The Fast-Minor flavor is constructed by replacing the major-scale notes with the corresponding minor-scale notes. The 'Slow Minor' flavor is constructed by setting the tempo of the 'Fast-Minor' flavor to $\tau = 0.5$ times the original tempo. For each minor flavor, the corresponding transient-flavor is constructed according to Section II. The pitch-gradation quantity, ρ varies between 0.5 and 0.7. In Fig. 4, the Slow Minor, Fast Minor and Fast Major flavors are marked by red colored crosses and the transient flavors are marked by red circles.

The listening experiment consisted of listening to and rating the tunes. First, the Slow Minor flavor of a tune was played and the participant had to choose the emotion that they associated with it. The list of emotions provided were the eight basic emotions in Plutchik's model [22]. However, the result is interpreted in terms of positive and negative valence rather than the exact emotions. Let a participant's response be E_1. That is, they associated the emotion E_1 with the Slow Minor flavor. Next, the participant listened to the Slow transient flavor and had to rate whether the association with E_1 increased, decreased, or remained the same (i.e. no difference). This was repeated for the other tune. The order of the two tunes was randomized across participants.

Next, the participants were presented with the Fast Minor flavor and they had to choose the emotion E_2 they associated best with it. Then, they listened to the Fast transient flavor and opined whether its association with E_2 had increased, decreased, or showed no difference. Thirdly, the participant

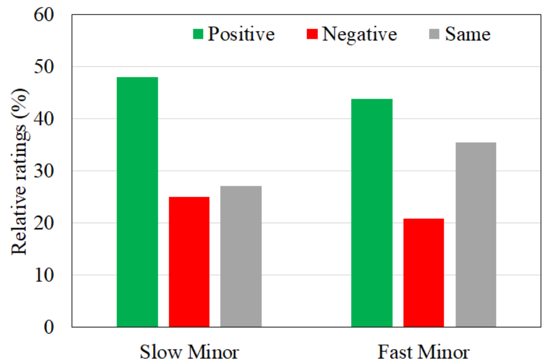

Fig. 5. Emotional valence of transients relative to the reference

chose the emotion E_3 that they associated most with the Fast Major version. Finally, they listened to the Fast transient flavor and opined whether its association with E_3 had increased, decreased, or showed no difference.

The survey described in the previous section was approved by our organization's Internal Review Board and was completed by 24 participants (15 male, 9 female; age range: 25-45 years) who gave informed consent. They hailed from different Indian states, had similar educational background with basic exposure to music terminology, and varied preferences for music-types. The number of responses for the emotions of the Slow Minor, Fast Minor and Fast Major flavors are categorized as 'Positive' and 'Negative', and given in Table I. It is encouraging that the Fast Major flavor, i.e. the melody tracks of the original tunes, is perceived as happy in 46 out of 48 ratings. This result matches the hypothesis of [16] in a completely different demography. Six out of 48 ratings associate a negative emotion when minor-scale notes are introduced and 22 ratings when the tempo is halved. The effect of transients is given in Fig. 5. In this case,

TABLE I

PERCEIVED AFFECT (VALENCE) OF THE TUNES

Flavor	Positive			Negative		
	Tune 1	**Tune 2**	**%**	**Tune 1**	**Tune 2**	**%**
Slow Minor	13	13	54	11	11	46
Fast Minor	22	20	87	2	4	13
Fast Major	24	22	96	0	2	4

'Positive' means that either (a) the emotion associated with the reference flavor is positive and the transient is perceived to have increased it or (b) the emotion associated with the reference flavor is negative and the transient is perceived to have decreased it. 'Negative' has the complementary meaning of 'Positive'. If the participant chose 'no difference', it is counted under the 'Same' category. The observations are:

1) The emotional impact of transients is felt more at slow tempos than at fast tempos.
2) When this impact is felt, the tunes with transients are approximately twice as likely to be perceived as more positive than as more negative, irrespective of the tempo.

IV. CONCLUSION

We propose a technique to generate a gradation of tunes in the two dimensions of time and pitch between an initial-reference tune and a target-reference tune. For audio-based intervention to influence affect, the target-reference tune is chosen as a tune considered happy (i.e., positive valence) by a majority of listeners. This perception was confirmed in our evaluation. While the proposed technique enables generation of many tunes, the perceptual evaluation was limited to the extremes. It showed that the impact of transients is felt more at lower than higher tempos. When felt, transients are twice as likely to be associated with an increase in valence, irrespective of tempo. Thus, rather than expecting the user to jump to affective states, the gradation enables designing audio intervention to progressively nudge the user towards a positive valence state while keeping them engaged throughout.

REFERENCES

[1] D. Kahneman and A. Deaton, "High income improves evaluation of life but not emotional well-being," *Proceedings of the national academy of sciences*, vol. 107, no. 38, pp. 16 489–16 493, 2010.
[2] Gavas et al, "A sensor-enabled digital trier social stress test in an enterprise context," in *EMBC*. IEEE, 2019, pp. 1321–1325.
[3] Ahuja et al, "Effect of music listening on p300 event-related potential in patients with schizophrenia: A pilot study," *Schizophrenia research*, vol. 216, pp. 85–96, 2020.
[4] E. Labbé, N. Schmidt, J. Babin, and M. Pharr, "Coping with stress: the effectiveness of different types of music," *Applied psychophysiology and biofeedback*, vol. 32, no. 3-4, pp. 163–168, 2007.
[5] Pallesen et al, "Emotion processing of major, minor, and dissonant chords," *Annals of the New York Academy of Sciences*, vol. 1060, no. 1, pp. 450–453, 2005.
[6] Viraraghavan et al, "A component-based approach to study the effect of indian music on emotions," in *Proc. Workshop on Speech, Music and Mind 2018*, 2018, pp. 31–35.
[7] ——, "Data-driven measurement of precision of components of pitch curves in carnatic music," *The Journal of the Acoustical Society of America*, vol. 147, no. 5, pp. 3657–3666, 2020.
[8] Krishna Raj et al, "An approach to transcription of varnams in carnatic music using hidden markov models," in *Proc. of the National Conference on Communication (NCC)*, 2017.
[9] David and Tan, "Emotional connotations of diatonic modes," *Music Perception: An Interdisciplinary Journal*, vol. 30, no. 3, pp. 237–257, 2012.
[10] Ali and Zehra, "Songs and emotions: are lyrics and melodies equal partners?" *Psychology of music*, vol. 34, no. 4, pp. 511–534, 2006.
[11] Balkwill, Thompson, and RIE, "Recognition of emotion in japanese, western, and hindustani music by japanese listeners 1," *Japanese Psychological Research*, vol. 46, no. 4, pp. 337–349, 2004.
[12] Kawakami et al, "Sad music induces pleasant emotion," *Frontiers in psychology*, vol. 4, p. 311, 2013.
[13] C.-F. Huang and S.-W. Wu, "The relationship between music excerpts and emotional responses of undergraduate students," *International Journal of Arts Education*, vol. 5, no. 1, pp. 33–53, 2007.
[14] Mathur et al, "Emotional responses to hindustani raga music: the role of musical structure," *Frontiers in psychology*, vol. 6, 2015.
[15] Hegde et al, "Variations in emotional experience during phases of elaboration of north indian raga performance," in *12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the Cognitive Sciences of Music*, 2012, pp. 412–413.
[16] Barbiere et al, "The color of music: Correspondence through emotion," *Empirical studies of the arts*, vol. 25, no. 2, pp. 193–208, 2007.
[17] V. S. Viraraghavan, A. Pal, R. Aravind, and H. Murthy, "Non-uniform time-scaling of carnatic music transients," *arXiv preprint arXiv:1711.02318*, 2017.
[18] Viraraghavan et al, "Visualizing carnatic music as projectile motion in a uniform gravitational field," in *Proc. Workshop on Speech, Music and Mind 2019*, 2019, pp. 31–35.
[19] ——, "Grammar-constrained music generation for unconstrained doodling," in *SMM*, 2020, pp. 6–10.
[20] "MIDI," https://www.midi.org/, accessed: 2020-08-17.
[21] Viraraghavan et al, "State-based transcription of components of carnatic music," in *ICASSP*. IEEE, 2020, pp. 811–815.
[22] Plutchik and Conte, *Circumplex models of personality and emotions.* American Psychological Association, 1997.