2025 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW) | 979-8-3315-1931-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICASSPW65056.2025.11011166

# Thoughts on Mapping and Interface Design of a Keyboard to Perform Continuous Pitch Ornamentations in Hindustani Music

Ninad Puranik*, Travis West†, Marcelo M. Wanderley†, Gary Scavone*

*Computational Acoustic Modeling Laboratory (CAML) †Input Devices and Music Interaction Laboratory (IDMIL)
Schulich School of Music, McGill University, CIRMMT, Montreal, Canada
ninad.puranik@mail.mcgill.ca travis.west@mail.mcgill.ca marcelo.wanderley@mcgill.ca gary.scavone@mcgill.ca

***Abstract*—The harmonium, a free-reed keyboard instrument that originated in Europe, is now widely used for vocal accompaniment in Hindustani music. However, it lacks the ability to produce the characteristic pitch ornamentations of Hindustani music, presenting several limitations to its use. In this paper, we present our ideas about developing a keyboard interface and mapping scheme for a digital harmonium that can be used to perform such pitch ornamentations and overcome the limitations of an acoustic harmonium.**

***Index Terms*—music technology, digital music instruments, music acoustics, free-reed instruments, harmonium, Hindustani music.**

## I. Introduction

The harmonium, invented in the early nineteenth century in Europe was introduced to the Indian subcontinent by the Portuguese and British settlers [1]. Being a European keyboard instrument, it uses a 12-tone equal temperament tuning. The instrument employs bellows-driven metallic free reeds for its sound production. The pitch of a free reed is mainly dependent on its length, which cannot be altered during a performance. Hence, a free reed instrument is incapable of performing significant pitch bending ornamentations.

While the harmonium grew in popularity in India, it also attracted severe criticism for its use of western equal-temperament tuning and its inability to perform the continuous pitch ornamentations such as 'meend', 'gamak' and 'andolan', which are a characteristic of North Indian (Hindustani) music [2].

While the instrument is unable to successfully render the pitch ornamentations, expert harmonium players have developed certain performance techniques to perceptually suggest them, and this has allowed the instrument to gain acceptance for use in vocal accompaniment in Hindustani classical music. Solo harmonium performances of Hindustani music are, however, still quite rare, in part because of this limitation.

To address the limitations of an acoustic harmonium, the authors were motivated to conduct research on creating a digital harmonium that can produce realistic harmonium sounds while allowing for the realization of the pitch ornamentations. The problem has been approached as a two-step process: i) to develop a natural sounding real-time capable synthesis model for harmonium sounds, and ii) to design an interface and mapping strategy to control the synthesis model.

While sample-based synthesis is a reliable method for realistic sound synthesis, it is quite a tedious process to record a large library of sampled sounds that includes all the nuances of an acoustic harmonium and map the performance gestures with the samples. Also, rendering the pitch bending ornamentations may require manipulation of the sampled sounds, which can lead to unwanted artifacts. In contrast, physics-based simulation models hold the promise of synthesizing natural and life-like sounds, while simultaneously allowing fine control over the synthesis parameters such as pitch [3]. Synthesis models for the harmonium have been previously researched by some of the authors [4]–[7], and an acceptable real-time solution has been demonstrated through a prototype of the digital harmonium [8].

This paper, therefore, will focus on the the second-step, i.e. the aspects related to the interface and mapping design for the novel instrument. We believe that harmonium players are more likely to adopt the new instrument in practice, if it can be played with their existing skills and they have perceived familiarity with the instrument. Hence, we aim to develop an instrument that enables pitch ornamentations, but is otherwise quite similar to an acoustic harmonium in feel, and in behavior.

## II. Performance techniques of Hindustani harmonium players

The interface and mapping of an acoustic harmonium is straightforward. The instrument uses a standard western keyboard of 7 white and 5 black keys in an octave. Models with a pitch range of 2.5 to 3.5 octaves are commonly used. A system of hand-pumped double-bellows is used to store pressurized air inside the instrument. The press of a key on the keyboard opens the passage for the pressurized air to pass around a set of reeds tuned to a specific note leading to the production of sound. A partially pressed key on the harmonium keyboard causes the air to flow with a lower rate around the reeds

The project has been funded by a student award from the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT).

leading to the note being played with a reduced loudness. Moreover, if a key is gradually pressed, it leads to the note being faded-in instead of having a sharp percussive onset. This is an important contrast to a piano where the key must be pressed fully for the internal hammer to strike the stretched strings and make a sound. Note expression and loudness in a piano thus depends on the velocity of the keys at the attack, while in a harmonium it depends on the evolution of the key-depression and the bellows pressure over time.

The overarching goal in melodic accompaniment for Hindustani music is to shadow the lead artist, usually the vocalist [9], [10]. In solo presentations of harmonium as well, harmonium players quite often perform the compositions from the vocal music repertoire and employ phrasings similar to a vocalist. It is therefore important to understand the harmonium player's intentions and techniques in the context of mimicking the musical phrases sung by a vocalist, while attempting to mask the 'deficiencies' in their instrument.

There are two primary kinds of human vocalizations used in singing, namely, the consonants and the vowels. While the vowel parts can be sung for long durations, and with pitch ornamentations, the voice necessarily breaks at the consonants. Any break in voice or an abrupt timbral change can be considered equivalent to a consonant, even if it does not correspond to an actual consonant in the sung lyrics. Expert harmonium players ensure that they perform, with a precise timing, the consonants in staccato fashion and the vowels in a legato fashion. To suggest a continuous pitch movement from one note to another during a vowel phase, the players fade out the previous note and fade in the next note by gently releasing one key while gradually pressing the next one. There is usually some overlap in this process when both the notes are heard simultaneously. This technique can be seen in the smaller yellow box in Fig. 1. It is believed that the gestalt principle of closure is responsible for the listeners to perceive this sound as a continuous pitch transition, especially when it is presented just after the vocal movement.

During the strong attacks while singing certain consonants, the pitch of the vocalist is not well-resolved, but the listeners prominently perceive the dynamic expressions. Some harmonium players rapidly press and release keys other than the main note to emulate such strong attacks and highlight the discreteness of voice at the consonants. This strong emphasis at the consonants also serves to underscore the contrast between them and the vowels which gently fade in and out, augmenting the perception of pitch continuity in the vowels.

While the bellows have the primary function to control the loudness of the harmonium sound, it is well known among harmonium experts that bellows pressure slightly affects the pitch of the notes [11]. A higher pressure in the bellows results in a slight lowering of the pitch of a note while a release of bellows leads to the note bending slightly higher. Some harmonium players make use of this phenomenon to suggest certain pitch movements. However, the effects of bellowing techniques are very small and it is debatable whether they are significant enough to be perceived. Moreover, the bellowing techniques are not consistent across the experts from different stylistic schools of harmonium for a universal gestural pattern to be observed. Hence, we will chose to disregard the effect of bellowing in our mapping design for the digital harmonium.

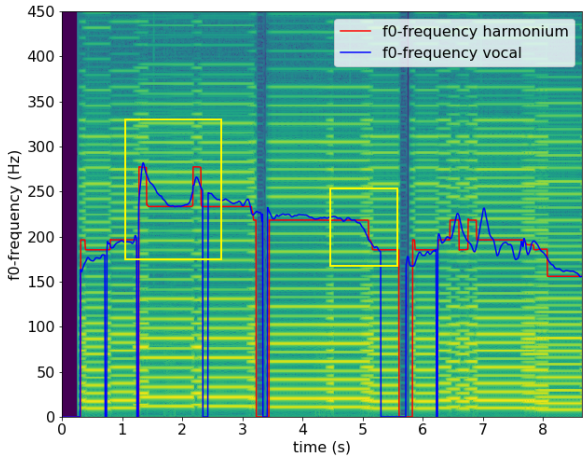


Fig. 1. A representative phrase in Raga Bhairav was performed vocally and on a harmonium by author 1, a Hindustani music and harmonium expert. Here, the estimated vocal and harmonium pitch contours are overlayed on the spectrogram of the harmonium sound sample. The yellow boxes highlight the harmonium player's approach of fading out the previous note and fading in the next note with a short overlap. The spectrogram corresponding to the vocal sound sample can be found in Fig. 3.

In a pitch contour, i.e. a plot of estimated pitch vs time, for a phrase sung by a vocalist (the blue curve in Figs. 1 and 3), there are some regions where the graph has large positive or negative slopes, while some regions where the plot is relatively flat. A pitch contour is also approximately flat for short durations near its local extrema and the points of inflection. Listeners perceive long flat regions in the pitch contour as the important notes in the melody, while regions with large slopes are considered as the transitions between these important notes. The short, approximately flat regions near the local extrema and the points of inflection are considered as notes that are 'touched' during a transition between one important note to another. These notes are known as 'kan-swars' which literally translates to 'granular notes'. Expert harmonium players ensure that they also give accurately timed gentle 'touches' of the 'kan-swars' in a pitch contour which enhances the illusion of a continuous pitch movement.

In summary, a harmonium player visualizes continuous pitch contours as short and long flat sections that are connected by transition zones. Playing the flat regions with accurate timing and rendering the transitions as a fading-out of the previous note and fading-in of the next note can then be considered as the basic approach followed by Hindustani harmonium players. This approach involves polyphony for the transition zones which is considered as improper by the purists. In the digital harmonium, we will attempt to remap the harmonium player's techniques to achieve a 'real' smooth

pitch transition instead of a mere 'perception' of seamlessness.

## III. Interface design

When a key is pressed on a piano, the mechanism of the instrument causes a felt-covered hammer to move towards the strings associated with that key. Past a certain point as the key is pushed down, the mechanism releases the hammer and further movement of the key up or down has no influence over the movement of the hammer. The velocity imparted into the hammer by the performer up to the moment when the hammer is released is the sole physical determinant of the force with which the hammer will strike the string. Thus, in performing an acoustic piano, hammer velocity (as determined by key velocity) is the principal determinant of variation in the sound resulting from any given key press. The MIDI protocol reflects the dominant piano-keyboard interface of commercial synthesizers at the time of its development [12]. Note events in MIDI correspond to piano key presses, and are associated with a pitch and velocity value. As in piano performance, variation of pitch and velocity are the main expressive dimensions of a performance of MIDI notes. Unfortunately, this conception of musical performance does not adequately represent the nuance of non-piano keyboard performance. In the case of the harmonium in particular, as described in Sec. II, various gestures are employed to mimic the performance of the vocalist. These gestures based on key depression are not generally captured by standard MIDI keyboards, which only measure the onset, velocity, and release of keys.

Various approaches have been employed to try to capture more information about keyboard performance gestures. Most prior work has been oriented towards enabling keyboard playing gestures that are not normally meaningful, such as wiggling a key from side to side, to be used for expressive control of sound parameters. Commercial products such as the Roli Seaboard [13] [1], Expressive E Osmose [2], and TouchKeys [14] are examples of this kind of approach. Osmose uses a proprietary keyboard that measures the depression and side to side movement of its keys to enable novel performance gestures. The Seaboard also uses a proprietary keyboard, made of a single foam surface with integrated pressure sensors measuring the cartesian position and pressure of touches on the surface. The TouchKeys employed capacitive touch sensing plates retrofitted to the top surface of a standard keyboard to capture novel performance gestures. Another example, in prior research, makes use of an infrared keyboard scanner that continuously measures the depression of the keys of a retrofitted keyboard [15]. This technology has been leveraged by Moro and colleagues in digital models of the Hammond organ as well as the flute [16].

Instruments like the Seaboard and Osmose can be viewed as departures from the traditional keyboard meant to evolve and expand the kinds of gestures that can be used by performers. There are benefits and drawbacks to this approach. These instruments naturally offer novel affordances to their players. In exchange, these players must learn to exploit these new affordances, and explore the new creative terrain offered by them. These instruments are not meant to be digital pianos. The advantage of a digital piano, in contrast, is that it directly serves piano players. For many musicians, who may not be interested in exploring novel musical and sonic terrain nor in learning new performance gestures, a digital piano is appropriate and advantageous. Such an instrument enables direct and faithful use of a piano players' existing skillset, as well as potential advantages over acoustic instruments, such as better portability, lower maintenance needs, and the option to access a variety of piano and keyboard instrument sounds. The approach of TouchKeys and Moro et al's keyboard scanner lies somewhere between these two extremes; since these technologies are retrofitted to existing keyboards, they unobtrusively integrate with performers' existing gestural vocabulary, while potentially opening up additional gestural modalities.

Our approach is somewhere in between the broad directions of augment or recreate, airing more similar to the development of a digital piano. We aim to capture the established gestures of harmonium performers, so that they can directly leverage their existing skillset, while offering the advantage of enabling them to more faithfully approximate the vocal performances that harmonium playing typically emulates. In order to accomplish this, we are developing an implementation based on Moro et al's infrared sensing schematic [15]. We will retrofit an existing harmonium keyboard with infrared sensors so as to continuously measure the displacement of the keys. This will enable the full nuance of harmonium players' gestural vocabulary to be captured, and mapped to sound parameters. Conveniently, most harmoniums can be easily disassembled, allowing the sensors to be installed inside the harmonium, underneath the keys. Installing the sensors underneath the keys will make them less obtrusive for the performers than placing them above the keys as done in prior work [15]. Whereas the top surfaces of the white and black keys of the keyboard have different heights, sizes and shapes, underneath the keyboard each key has uniform spacing and dimensions, simplifying the layout of the sensors. As well, operating the infrared sensors in a consistent dark environment will improve the signal-to-noise ratio of the measurements by reducing exposure to ambient light.

## IV. Mapping

Given an accurate and reliable measurement of the harmonium player's gestures, a strategy must be devised to 'use' these gestures to achieve their 'intended' effect in the synthesized sound. We present here some ideas about mapping of these gestures to achieve the sound outcome desired by the harmonium player.

For consistency with an acoustic harmonium, when a single key is pressed on the keyboard, the corresponding note should be synthesized. The amplitude of the synthesized note should depend on the depth of key press, with the highest value for a fully pressed key. If a player presses two or more

[1] https://roli.com/eu/product/seaboard-2

[2] https://www.expressivee.com/2-osmose

keys simultaneously, we should assume that a polyphony is desired. In such cases, all the corresponding notes should be synthesized simultaneously, with their amplitudes dependent on their respective key depressions.

Effectively, a remapping is only required with respect to an acoustic harmonium, for the case when two keys are played in succession in a legato fashion, i.e. when one key is released while another is being pressed. As described in Sec. II, the intended effect of this gesture is to perform a continuous glide ('meend') between the two notes in consideration. There can be an infinite ways to join two points on a pitch contour with a continuous curve. However, we will assume that the join is made by a monotonic and differentiable 'S-curve'. The assumption of monotonicity is justified since the harmonium player is expected to assume any peak, valley or a point of inflection as an 'approximately flat' zone and press a key to highlight it. The differentiability condition has been imposed since we expect lesser synthesis artifacts when the pitch is varied as per a smooth curve.

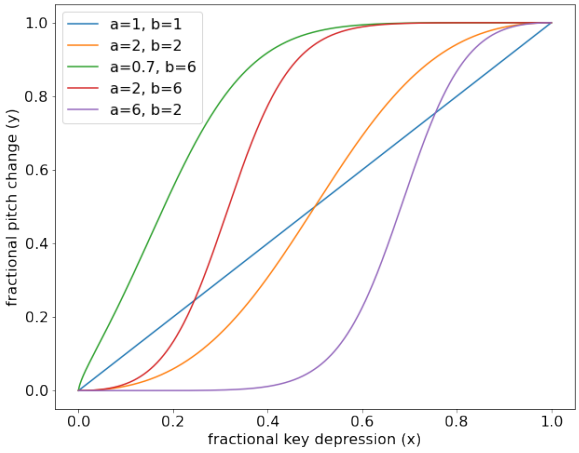


Fig. 2. Visualizing S-curves with different parameters. $a > 1, b > 1$ is required for the S-curves to be flat at the beginning and end of a transition.

We propose to use the following family of S-curves to approximate the pitch contours in the synthesis:

$$y = \frac{(\Delta x)^a}{(\Delta x)^a + (1 - \Delta x)^b}, \qquad \Delta x = \frac{x_2}{x_1 + x_2}$$

where $\Delta x$ is the fractional key depression related to $x_1$ and $x_2$, the key depression of the previous note and the next note respectively. $y$ represents the fractional change in pitch. The synthesized pitch can be obtained from $y$ by the expression:

$$synth\_pitch = old\_pitch + y \times (new\_pitch - old\_pitch)$$

With this definition, $0 \leq x, y \leq 1$. Parameters $a$ and $b$ control the flatness of the S-curve at the beginning and the end of the transition respectively as can be seen in Fig. 2.

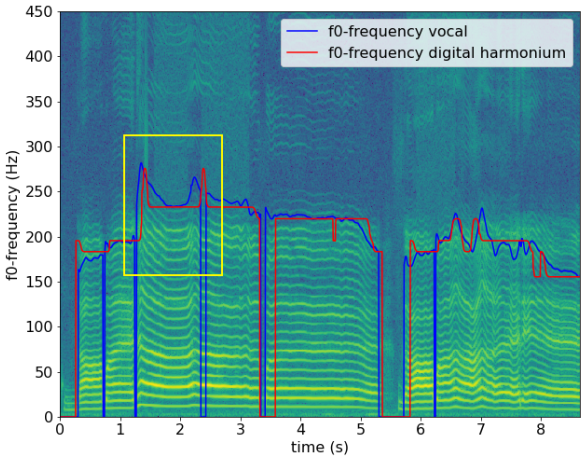


Fig. 3. A comparison of pitch contours for the same melodic phrase performed by a vocalist and a digital harmonium with MIDI mapping. The pitch contours are overlayed on the spectrogram of the vocal sample. The relative positions of the red and blue peaks in the yellow box highlight the timing errors and show that the slow glides were not rendered with sufficient accuracy.

## V. Results and Conclusions

Since our customized interface with the infrared sensors is yet to be built, we present the results with our approach adapted for a standard MIDI keyboard and the synthesis implemented as Pure-Data patch. In the modified mapping, when a new 'note-on' message is received before a 'note-off' message of the previous note, we assume that the value of x varies linearly from 0 to 1 in a time duration inversely related to the MIDI velocity parameter of the next note. Parameters 'a' and 'b' were assumed to have the values 2 and 6 respectively.

It can be seen in Fig. 3 that the red curve corresponding to the pitch contour with the modified mapping has gradual and rounded transitions between the flat regions. However, a considerable variation with respect to the vocal pitch contour (blue curve) is apparent. The performer (author 1) experienced difficulty to control the pitches with the new instrument. Since a glide could start only after the note-on message for the next note was received, there was a significant lag between the gestures and the sound, causing timing inaccuracies in the performance. The maximum duration of the glide had to be limited due to the delay, making it impossible to render the slow glides with sufficient accuracy.

While our results with the MIDI based mapping are encouraging, they also underline the necessity of developing a customized interface as proposed in Sec. III to capture the gestural nuances. Domain experts may have some disagreement with the simplifying assumptions about the harmonium players' gestures that we have made in Sec. II for the ease of engineering. Hence, future work also involves taking feedback from harmonium experts from different stylistic schools to improve the playability of the novel instrument while working within the engineering constraints.

## REFERENCES


[1] B. Abels, *The Harmonium in North Indian Music*. New Age Books, 2010.

[2] M. Rahaim, "That Ban(e) of Indian Music: Hearing Politics in The Harmonium," *The Journal of Asian Studies*, vol. 70, no. 3, pp. 657–682, Aug. 2011. [Online]. Available: https://www.cambridge.org/core/product/identifier/S0021911811000854/type/journal_article

[3] J. O. Smith, *Physical Audio Signal Processing*, 10th ed. http://ccrma.stanford.edu/~jos/pasp/, online book, 2010.

[4] N. V. Puranik and G. P. Scavone, "Physical modelling synthesis of a harmonium," in *Proceedings of Meetings on Acoustics, Fourth Vienna Talk on Music Acoustics*, vol. 49, no. 1. Acoustical Society of America, 2022, p. 035015.

[5] N. Puranik and G. Scavone, "Physically inspired signal model for harmonium sound synthesis," in *Proceedings of the 26th International Conference on Digital Audio Effects*, 2023.

[6] ——, "Clamped bar model for free reeds," in *10th Convention of the European Acoustics Association, Forum Acusticum*, 2023.

[7] N. V. Puranik and G. Scavone, "Aeroacoustic modeling of blown-closed free reeds," *The Journal of the Acoustical Society of America*, vol. 155, no. 3_Supplement, pp. A196–A196, 03 2024. [Online]. Available: https://doi.org/10.1121/10.0027284

[8] N. V. Puranik, "Harmonium - ARTASIA," August 2024, live lecture-demonstration concert in the 7th edition of Arts appreciation seris - 'Artasia', organized by Centre-Kabir, Montreal, Canada. [Online]. Available: https://www.centrekabir.com/wp-content/uploads/2024/07/ArtAsia-Poster-2024_EN_mod_20240624_113121_0000.jpg

[9] J. J. Napier, "The introduction and use of the harmonium in north indian classical music," Master's thesis, UNSW Sydney, 1994.

[10] ——, "Sangat: melodic accompaniment in north indian vocal music," Ph.D. dissertation, UNSW Sydney, 2001.

[11] A. Thatte, *Sangeet Vimarsh*. Swanandi Prakashan, Pune, 2010.

[12] G. Loy, "Musicians make a standard: The MIDI phenomenon," *Computer Music Journal*, vol. 9, no. 4, pp. 8–26, 1985.

[13] R. Lamb and A. Robertson, "Seaboard: A new piano keyboard-related interface combining discrete and continuous control." in *Proceedings of the International Conference on New Interface for Musical Expression*, 2011.

[14] A. P. McPherson, "TouchKeys: Capacitive multi-touch sensing on a physical keyboard." in *Proceedings of the International Conference on New Interface for Musical Expression*, 2012.

[15] G. Moro, A. McPherson *et al.*, "A platform for low-latency continuous keyboard sensing and sound generation," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2020.

[16] G. Moro, "Beyond key velocity: continuous sensing for expressive control on the hammond organ and digital keyboards," Ph.D. dissertation, Queen Mary University of London, 2020.