IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 15, 2020, accepted July 25, 2020, date of publication August 28, 2020, date of current version September 21, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3019978

Hands-Free Accessible Digital Musical Instruments: Conceptual Framework, Challenges, and Perspectives

NICOLA DAVANZO AND FEDERICO AVANZINI
Department of Computer Science, University of Milan, 20133 Milan, Italy
Corresponding author: Nicola Davanzo (nicola.davanzo@unimi.it)

ABSTRACT Exponential increases of available computational resources, miniaturization, and sensors, are enabling the development of digital musical instruments that use non-conventional interaction paradigms and interfaces. This scenario opens up new opportunities and challenges in the creation of accessible instruments to include persons with disabilities into music practice. This work focuses in particular on instruments dedicated to people who can not use limbs, for whom the only means for musical expression are the voice and a small number of traditional instruments. First, a modular and adaptable conceptual framework is discussed for the design of accessible digital musical instruments targeted at performers with motor impairments. Physical interaction channels available from the neck upwards (head, mouth, eyes, brain) are analyzed in terms of potential and limitations for musical interaction. Second, a systematic survey of previously developed instruments is presented: each is analyzed in terms of design choices, physical interaction channels and related sensors, mapping strategies, performer interface and feedback. As a result of this survey, several open research directions are discussed, including the use of unconventional interaction channels, musical control mappings, multisensory feedback, design, evaluation, and adaptation.

INDEX TERMS Accessibile interfaces, music technology, digital musical instruments, inclusive music practice.

I. INTRODUCTION

Music playing is one of the most universally accessible and inclusive human activities and is part of all known cultures [1]. Its engaging power applies to all ages [2], and is known to provide benefits also in terms of non-musical skills [3], [4]. Yet, music playing is still not easily accessible for persons with disabilities.

Digital musical instruments (DMIs hereafter) [5] have the potential for augmented accessibility, as they allow for new, non-conventional modes of interaction. These are instruments in which sound generation is based on digital means and is achieved by the performer through physical actions detected by sensing devices. Thanks to the exponential increase of computational power, miniaturization, and available sensors, research on DMIs has expanded during the last two decades into the use of innovative interaction paradigms and interfaces. Although the majority of commercial DMIs still uses the piano keyboard as the main interface, over the years the research community has explored different physical channels and sensors [6]. Nonetheless, the main employed physical channels remain the hands, the feet, and breath.

The term "accessible DMIs" (ADMIs) refers to instruments designed for persons with disabilities. A distinction can be drawn between "performance-focused" and "therapeutic" instruments [7], where the former include ADMIs designed to enable masterful performances by musicians with disabilities, while the latter include instruments designed to elicit therapeutic or wellbeing aspects of music making, even for non-musicians.

Some recent works have provided broad surveys of ADMIs, including both research projects and commercial products. Frid [8] provides an extensive review which encompasses several target groups, including persons with physical, sensory, and cognitive impairments, persons with complex needs or special educational needs, elderly or young children. The review of Larsen *et al.* [9] has a similar broad focus. Graham-Knight and Tzanetakis [10] propose a set of

The associate editor coordinating the review of this manuscript and approving it for publication was Cristina Rottondi.

principles for how to work with people with disabilities to develop a new musical instrument.

A large percentage of the ADMIs reviewed by Frid [8] focuses on user groups with physical impairments (39.8%). However, many of these are aimed at musicians with partially able limbs [11]. On the other hand, only 4.8% of the reviewed ADMIs are devoted to persons with quadriplegia. This work focuses precisely on instruments dedicated to people who can not use upper and lower limbs. For these persons, the only acoustic means for musical expression are the voice and a limited number of instruments, such as the whistle, the mouth harmonica, the kazoo, and possibly a few more. We will use the term *HeaDMI*s to refer to DMIs that only use physical interaction channels on the head, from the neck upward.

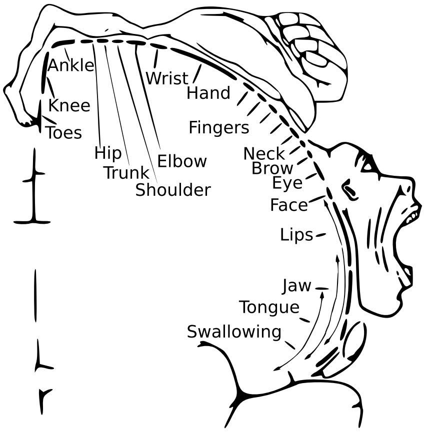

FIGURE 1. **The motor homunculus, a topographic representation of body areas on the motor cortex [12].**

The fact that some groups of muscles can be controlled more finely than others has well-known implications for the design of human-computer interfaces [13]. One determinant – albeit not the only one – of the performance of a group of muscles is the portion of the motor cortex devoted to it. Fig. 1[1] shows the representation of the motor homunculus obtained by Penfield and colleagues [12] from the mapping of the motor cortex: one relevant aspect is that physical channels located on the head occupy the second largest area of the motor cortex, after the hand and fingers (which reflects their evolutionary importance for verbal and non-verbal communication). This provides support to the idea of using these channels (plus the brain cortex itself) as musical controllers [14].

The potential target population for HeaDMIs is vast. Sears *et al.* [15] presents an overview of health conditions and related physical impairments that affect the upper body and consequently hinder the use of traditional computer interfaces. These include diseases and congenital disorders such as amyotrophic lateral sclerosis, multiple sclerosis, muscular distrophy, transverse myelitis, amelia, as well as traumas and injuries, such as limb amputations, stroke, spinal cord injuries. Many of these can lead to locked-in syndrome, a condition in which a person is awake and conscious but can only communicate through the eyes. The size of the affected population can be inferred by epidemiological data. As an example, the incidence (occurrence of new cases) of spinal cord injuries varies from developed countries (13.1-163.4 cases per million) to undeveloped countries (13.0-220.0 cases per million) [16], with 250,000 to 500,000 persons affected every year worldwide [17]. Rates of prevalence (persons affected at a given time) range from 906 per million in the US (highest recorded) to 250 per million in France (lowest recorded) [18]. As a further example, the incidence of amyotrophic lateral sclerosis is 1.0-2.6 cases per 100,000 people every year [19], and is particularly high in Europe, with 15,000 new cases per year [20], [21]. Prevalence ranges from 4 to 9 people per 100,000 [19]–[21].

In order to be able to create a truly accessible instrument, we should redefine the concept of ADMI as one modeled entirely on the residual motor abilities of the person, whose interaction parameters adapt to the best available physical channels. This is the main motivating concept for the present work, where we aim at developing a modular and adaptable conceptual framework for HeaDMI design, by which various types and levels of physical impairments can be addressed. First (Sec. II) we propose a general structural diagram for a HeaDMI, and we use it to revisit a set of DMI-related relevant definitions and design issues in the light of our specific focus. Building on this, we compile (Sec. III) an exhaustive list of all the physical interaction channels that are available from the neck upwards: for each channel we review its uses for human-computer and musical interactions, and we discuss their potential and limitations with respect to a set of relevant channel properties. We then provide (Sec. IV) a systematic survey of previously developed HeaDMIs, including research projects and commercial products, many of which were not included in previous reviews on generic ADMIs. To this end, each reviewed instrument is analyzed along a set of technology-, design-, and music-related dimensions (this approach is partly inspired by a previous systematic survey on sonification strategies [22]). This survey provides the ground for discussing (Sec. V) open challenges and future research directions.

II. DMIs REVISITED

The term DMI has somewhat fuzzy boundaries in the literature, and intersects with other forms of musical interfaces: we adopt a rather restrictive definition based on Malloch *et al.*'s conceptual framework [23], which is in turn inspired by Rasmussen's model of human information processing [24]. In this view, interaction behaviors can be skill-, rule-, or model-based. Briefly, skill-based behaviors are related to activities which take place without conscious attention as smooth, automated, and highly integrated movements controlled on the basis of continuous signals coming from the environment: playing a conventional acoustic instrument

---

[1] Adapted from a figure licensed under CC-BY-SA-4.0 license: https://commons.wikimedia.org/wiki/File:Motor_homunculus.svg

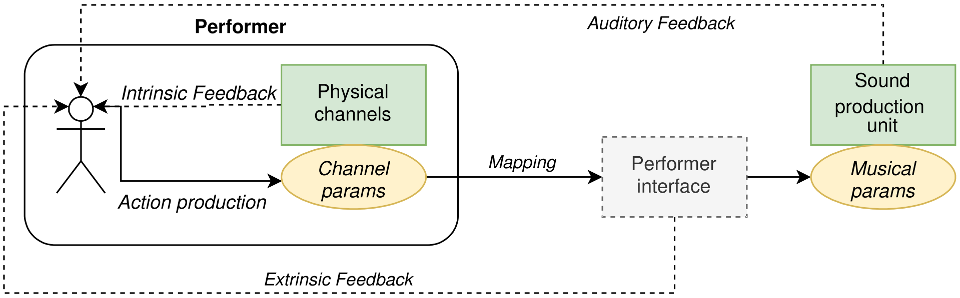

FIGURE 2. Structural diagram of a HeaDMI.

falls in this category, along with hand-writing, sports, bicycle riding, etc. In rule-based behaviors, activities consist of subroutines controlled by stored rules or procedures which have been learned or derived empirically, and information from the environment is typically perceived as signs: musical examples in this category include sequencing, live diffusion, creating a rhythm on a drum machine, etc. Model-based behaviors refer to more abstract activities in which performance is directed towards a conceptual goal (algorithmic music composition, presentation of recorded material, etc.) and information is perceived as symbols.

The above discussion provides the ground for stating that in this work we focus on skill-based musical instruments, which bear close similarities to traditional ones in terms of both performance behavior and context. It has however to be noted that the distinction between skill- and rule-based behaviors is generally blurred [24], and depends on previous training and experience. This is true for musical performance as well [23], where rule- and skill-based behaviors are mixed.

Fig. 2 presents a structural diagram for a HeaDMI, whose components are discussed next. Similar diagrams have been proposed for DMIs. Miranda and Wanderley [5, Fig. 1.1, p.3] focused on the separation between sound production unit and control. McGlynn [25, Ch.3] introduced the performer into the diagram. Here we further specify the performer's action in terms of physical interaction channels, parameters, and mapping strategies, as well as the sources of feedback to the performer.

A. PHYSICAL CHANNELS

The physical channels used to interact with traditional musical instruments are hands, fingers, breath, mouth/lips, and feet, with rare exceptions. A HeaDMI can only exploit the remaining able channels of the performer: these may therefore include head movements, gaze pointing, mouth aperture, etc. A comprehensive list and analysis of these channels is presented in Sec. III.

Any single physical channel can have multiple associated parameters: as an example, those associated to *Head movements* include pitch, yaw and roll angles, while those associated to *Gaze pointing* are the 2D pointing coordinates on the screen, the duration of fixations, etc. Each parameter can be estimated by appropriate sensors and can be assigned a role in the musical interaction.

B. SOUND PRODUCTION UNIT

This block is responsible for the sound synthesis. The possibility of separating the control interface from the sound production unit in a DMI provides an additional degree of freedom with respect to acoustic instruments. The sound production unit generally exposes an interface which is able to receive a set of messages and events influencing musical parameters.

We opt for an operative classification of musical parameters, often used in the context of DMIs [26], [27], which identifies three levels of control over musical processes: the *Note* level requires parameters related to a single note event; the *Timbral* level demands parameters with high temporal resolution, acting on timbral sound properties, even within a single note event; the *Process* level is a "macroscopic" one, which is associated to global or structural musical characteristics. These three levels are reported in Table 1, along with a non-exhaustive list of possible associated musical parameters (more precisely, the parameters listed in Table 1 are those used by the HeaDMIs reviewed in Sec. IV).

TABLE 1. A non-exhaustive list of musical parameters, associated to different control levels.

Control Level	Musical parameters	Abbr.
Note	Note on/off	note
	Pitch	pitc
	Intensity	inte
	Glide	glid
	...	
Timbral	Vibrato	vibr
	Brightness	brig
	Sustain	sust
	...	
Process	Instrument selection	isel
	Mode/scale selection	msel
	Transposition	tran
	Harmonic change	hcha
	Tempo	temp
	Panning	pann
	...	

At the Note level, note on/off events control the triggering/releasing of a note; pitch refers to the perceived note height, and may be quantized on a musical scale or may vary continuously; intensity relates to the energy injected into the note emission, and thus to its loudness but also to its spectral coloration; glide refers to a smooth transition between pitches of two successive notes.

At the Timbral level, vibrato is a rapid, slight oscillation in pitch which produces a richer tone; brightness refers to the possibility of manipulating the spectral energy of the sound towards the high or the low frequencies; sustain is a control available on some instruments (e.g., on the piano through the damper pedal) by which the note keeps resonating after its actual release (possibly along with sympathetic resonance from other notes).

At the Process level, instrument selection refers to the used instrumental sound; mode/scale selection and transposition refer to the possibility of redefining the musical scale on which the instrument is tuned or to transpose all the pitches by a given offset (e.g., an octave); harmonic change controls the switching between different chords; tempo refers to the control over the beats per minute (BPMs) of the music being played; panning controls how the sound is distributed on output channels (e.g., in a stereo or multichannel set-up).

C. MAPPING

By mapping we refer to the way in which channels parameters are linked to musical ones. McGlynn [25] discusses various mapping strategies, some of which are especially relevant for the channels analyzed in Sec. III.

Trigger: an action of the physical channel causes an instantaneous event (an example is a hit on a drum pad). *Toggle:* an action causes an instantaneous switch to a different state, and a subsequent one returns it to the previous state (an example is the use of a selector to transpose by one octave up and down, or to switch from one scale to another). *Counter:* different actions allow to scroll between different states in a circular fashion (as an example, pressing a key on an electronic keyboard allows to scroll through different available sounds). *Hold:* an action changes the internal state of the system, as long as it is maintained (an example is the pressure on the expression pedal of a piano). *Continuous range:* the value of a physical channel parameter in a continuous range is mapped to a musical parameter over an analogously continuous interval (as an example, breath pressure can be mapped to intensity, or head position can be mapped to pitch). *Discrete range:* the value of a physical channel parameter is quantized and mapped to a discrete set of values of a musical parameter (as an example, in a harmonica the horizontal head position is mapped to a discrete set of pitches). *Excitation:* the rate of change (time derivative) of a physical channel parameter is mapped to a musical parameter (as an example, in a violin the bow speed affects the intensity).

Mappings also have associated qualities, which depend on strategies but also on the physical and musical parameters involved, and have a major influence in instrument playability, expressiveness, and enjoyment. Some of these qualities are particularly relevant for HeaDMIs.

Transparency. This quality refers to the "psychophysiological distance" [28] between physical and musical parameters of the mapping, from both the performer and the audience perspective. For the former, transparency depends on cognitive understanding of the mapping and on the level of dexterity with the instrument, while the latter only need to have an understanding of causal relationships between performer's actions and sonic results. For both, understanding is derived from previous knowledge and expectations: as an example, mimicking physical actions on an acoustic instrument, or using metaphors (e.g., pitch increasing from left to right as in a piano keyboard), aids transparency. This aspect is particularly relevant for HeaDMIs, due to the unconventional physical channels considered.

Energy. One particularly important ecological principle (i.e., one reflecting expectations derived from everyday experience) is that the acoustic energy of the instrumental sound should be the product of muscular energy injected by the performer's gestures into the instrument. Usability experiments in DMI design [29] have shown that incorporating energy into the mapping provides a more engaging natural instrument and a tighter connection of the performer to it. The physical channels considered in this work allow for limited possibilities of movement and muscular activation. It is therefore necessary to maximize the use of energy in the mapping, and also to devise alternative strategies to compensate for these limitations.

Cardinality. Simple mappings employ one-to-one relationships between physical and musical parameters. However it has been long been suggested [29] that relationships involving higher cardinalities (many-to-one, one-to-many, many-to-many) should be preferred especially when several musical parameters are exposed by the sound production unit. These relationships, which are typical of most acoustic instruments, have been shown to be more rewarding and intuitive for musical interaction and to provide more expressive control, possibly at the expense of longer learning times [29]. They may also require additional layers of processing to extract intermediate parameters [30]: as an example, parameters from several physical channels may be combined to estimate the performer's facial expressions and control mode selection in a many-to-one mapping.

D. PERFORMER INTERFACE

As mentioned in Sec. II-B, in DMIs the instrument interface is often physically separable from the sound production unit. In HeaDMIs such interface may be totally absent, especially whenever the employed channels do not require external references for performing their actions. In this case the body of the performer becomes the interface to some extent. This would be the case for many of the channels discussed in Sec. III. Some other channels, like Gaze pointing, require visual objects on a screen in order for gaze fixations to occur and be detected. Several of the brain-computer interfaces

TABLE 2. List of physical channels analyzed in Sec. III, with related parameters and sensors. Refer to to Sec. III-A for sensors and channel properties. Property values are labeled as H (high) or L (low), X (yes). Asterisks in the Rest column indicate that the property applies to a subset of channel parameters only.

Group	Channel name	Abbr.	Main parameters	Sensors	Resolution	Fatigue	Involuntary	Stability	Persistence	Rest	Smoothness	Accuracy	Velocity
Eyes	*Gaze pointing*	gaze	2D coordinates, Fixation duration	eyet, cam	H	L	X	L	H		L	H	H
	Eye movements	eymv	2D angular displacements	eog, eyet, cam	H	L	X	L	H		L	H	H
	Blinking	blnk	Displacement (boolean), duration	eyet, emg, eeg, acc, cam	L	L	X	L	H	X	L	L	H
	Eyebrow movements	eybr	Vertical displacement	emg, eeg, acc, cam	L	H	X	H	L	X	L	L	H
Mouth	*Voice*	voic	Pitch, Intensity, Timbre, Speech	micr	H	L		H	L	X*	H	H	H
	Whistling	whis	Pitch, Intensity	micr	H	H		H	L	X*	H	H	H
	Breath	brea	Pressure, Airflow	brea, micr	H	H	X	L	L	X	H	H	H
	Mouth-lip movements	mout	Height, Width, Aspect ratio, Area	emg, acc, cam	H	L		H	H	X	H	L	H
	Tongue	tong	3D Tip coordinates, Pressure	hall, piez, btn, acc	H	L	X	H	H	X	H	H	H
	Teeth	teet	3D coordinates, Pressure	piez, btn	H	L	X	H	H	X*	H	L	L
Head	*Head movements*	head	6D coordinates	eyet, acc, head, dept, gyro	H	H		H	H		H	H	L
	Neck tension	nkte	Tension	emg	L	H		L	L	X	L	L	L
Brain	*Mental states*	ment	PSA, SCA, HJA, TAV	eeg	L	L	X	L	L		L	L	L
	Attention	atte	P300, SSVEP, ERNP	eeg	L	H	X	L	L		L	L	L
	Imagery	imag	ERS/ERD, SCP	eeg	L	H	X	L	L		L	L	L

discussed below also require external stimulation that elicits the desired brain responses.

If present, the user interface may be represented by physical or virtual objects (e.g. shown on a screen). Furthermore, it may be part of the mapping strategies of the instrument, as physical channel parameters may manipulate static or dynamic objects (keys, sliders, or more complex elements), and this manipulation would reflect on musical parameters.

This component can be also devoted to providing extrinsic feedback to the performer, in addition to the intrinsic (tactile, proprioceptive, kinesthetic) feedback generated by performer's movements [31] and the auditory feedback provided by the sound production unit. Extrinsic feedback may use several modalities (e.g., visual feedback through computer screen, vibrotactile feedback through actuators, etc.). In acoustic instruments, a physical interface has the double function of mapping actions to sound and of providing extrinsic feedback to the performer: as an example, the piano keyboard provides both the mechanical machinery that sets strings into vibration, and a visual and haptic interface for the performer to locate pitches and control the dynamics. In DMIs, rich, multimodal feedback can be introduced to enhance the interaction between player and instrument [5].

E. EVALUATION

Evaluation of DMIs encompasses a broader set of aspects than those typically considered in HCI. This is because evaluation spins around the concept of performance, which involves a number of stakeholders: performers, composers, audiences, designers, manufacturers. O'Modhrain [32] proposed a framework for DMI evaluation which takes into account the roles of stakeholders in a number of evaluation goals (enjoyment, playability, robustness, compliance to design specification).

For DMIs in general, the causal link between the performer's gestures and the sound generation may not always be clearly perceivable by the audience, and may impact negatively on the relationship between performers and audience. Augmenting the audience experience through additional sensory channels (e.g., visuals [33] or haptics [34]) can help reestablish such a link by increasing the transparency of the interaction.

This issue is even more challenging in the case of HeaDMIs, where mappings have necessarily a limited ecological validity, and transparency from the audience perspective is reduced by the lack of apparent exchange of energy between the performer and the instrument. In other words, the need for additional multimodal cues, able to reestablish the connection between cause and effect, becomes particularly important in order to make a HeaDMI's performance convincing and expressive.

III. ANALYSIS OF PHYSICAL CHANNELS

This section presents an analysis of the physical channels that can be exploited for designing interaction in HeaDMIs. For the sake of clarity, the analyzed channels are clustered into four groups, namely Eyes, Mouth, Head, and Brain. A summary is provided in Table 2.

A. CHANNEL CHARACTERIZATION

In the remainder of this section, each analyzed physical channel is characterized in terms of its associated parameters, along with (i) a list of suitable sensors for measuring such parameters, and (ii) a number of channel properties.

The properties considered here are visible in Table 2 (rightmost 9 columns). These were empirically selected on the basis of their utility to characterize a physical channel in terms of the mappings that it could be used for, with respect to both musical parameters and mapping strategies. *Resolution* refers to the number of distinct values that channel parameters can achieve, and relates to the possibility of performing fine motor actions, accessing for example a higher number of discrete pitch values. *Fatigue* indicates whether the prolonged use of a channel can easily tire the performer, requiring rest after a performance period. *Involuntary movements* can be present and interfere with voluntary use, given the nature of the channels. *Stability* indicates the possibility of maintaining a stable value (e.g., a given intensity or pitch) without flickering. *Rest* indicates whether the channel possesses a natural, stable and easily accessible rest state in which fatigue is minimized, which can be mapped to special values (e.g., zero intensity or no sound emission). *Persistence* indicates whether there is a physiological need to return to the rest state after a period of use, as in the case of breath emission. *Smoothness* indicates the degree of fluidity in changes of the channel parameters, as opposed to movement jerkiness, typical for example of gaze movements. *Accuracy* refers to the ability of hitting a target value with the smallest possible error, needed for example to move through pitch values. *Velocity* refers to the maximum attainable movement speed, regardless of accuracy, needed to move quickly through different parameter values.

A detailed characterization of the above properties would require to collect, analyze, and compare large amounts of physiological data. Moreover, it is known that user models calibrated exclusively on able-bodied subjects are not applicable to persons with various motor impairments [35]. In light of these considerations, for the purpose of this work we resort to a qualitative characterization in which these properties are only given binary values (high-low, yes-no) with reference to able-bodied persons, which serve as a best-case reference for motor-impaired persons.

TABLE 3. List of sensor names and abbreviations.

Sensor name	Abbr.
Accelerometer	acc
Breath sensor	brea
Button	btn
Camera	cam
Depth (IR) camera	dept
Elecroencephalographic headset	eeg
Electromyographic sensor	emg
Electrooculographic sensor	eog
Eye tracker	eyet
Gyroscope	gyro
Hall effect sensor	hall
Head tracker	head
Microphone	micr
Piezoelectric pressure sensor	piez

The considered sensors are summarized in Table 3. Many of these do not require further comments, while more complex ones merit additional discussion.

Eye trackers detect the 3D position of pupils and the absolute gaze point in the visual scene (e.g., on a screen). Majaranta and Bulling [36] provide a review of eye tracking technologies. In this article we mainly refer to non-invasive eye trackers equipped with infrared cameras, which require short calibration times and work in natural exposure conditions, some of which are now available at relatively low prices. An alternative sensing technology is electrooculography, in which eye movements are detected by measuring the standing corneal-retinal potential arising from hyperpolarizations and depolarizations. Electrooculographic sensors consist of electrodes (typically 5) placed on the skin around the eye. They have some advantages over video- and infrared-based tracking, namely independence on lighting conditions, lightweight signal processing, mobile implementations.

Electromyographic sensors detect muscle activations by measuring electrical activity. The potential of these sensors for the development of ''muscle-computer'' interaction has long been recognized [37].

Head trackers detect head rotation angles (pitch, yaw, roll) and possibly translatory degrees of freedom. Hess [38] provides a categorization of current head-tracking technologies. The 3D position of the eyes can be used also for head tracking purposes, and many commercial eye trackers possess this feature, thus enabling the use of two physical channels through one sensor.

Electroencephalographic (EEG) headsets consist of a cap with electrodes. This brings about the more general topic of Brain-Computer Interfaces (BCIs), which are well discussed in several surveys [39]–[41]. BCIs are operated by detecting brain signals through more or less invasive sensors, chiefly EEGs. Although alternative techniques exist (magnetoencephalography, functional Magnetic Resonance Imaging, functional Near-Infrared Spectroscopy), EEG sensors have several advantages in terms of invasiveness, costs, and temporal resolution. BCIs are often categorized as ''passive'' (using arbitrary brain activity without the purpose of voluntary control), ''reactive'' (using brain activity arising in reaction to external stimulation), and ''active'' (using brain activity that is consciously controlled by the user, independently from external events). The three brain physical channels discussed in Sec. III-E map one-to-one into these categories.

B. EYES

1) GAZE POINTING

Movements of the gaze point on an object or surface have some peculiar characteristics [36]. Saccades are rapid and short ($\sim$30 ms) movements. Involuntary saccades can be stimulated by fast or unexpected objects, and by the absence of reliable reference points. During fixations, the gaze point is still and focused on a narrow area. Fixations between two subsequent saccades have typical durations of $100 - 400$ ms. Involuntary angular jitter ($\sim 0.1^\circ$) can occur during a fixation.

Finally, smooth pursuits occur when the gaze follows a moving target.

Gaze pointing has well-established applications in HCI, including mouse emulation, gaze-based text entry, web browsing, gaze-controlled games, attention-aware interfaces, user modeling and monitoring [36]. Hornof [42] reviews eye-controlled music performance systems, which in most cases allow interaction with pre-defined compositions (by triggering samples and musical events), or control over music production software. He also proposes an interesting analysis of the capabilities and constraints of this channel in relation to musical expression.

One prominent challenge is the so called "Midas touch" problem [36]: fixations on an interface element may lead to its activation even when the user has no such intention. Moreover, elements crossed by saccadic movements may also be activated. Typical solutions include introducing a dwell-time (a short delay to detect fixations), using specially designed selection areas and gaze gestures, or exploiting a second physical channel to perform activations. For musical instruments, this problem can prevent the completion of basic actions (e.g., jumping between two non consecutive pitches).

The main parameters are the 2D gaze point coordinates on the screen, as well as fixation duration. Given its properties, gaze pointing is particularly suited for Continuous and Discrete range selection mapping strategies. Hold strategies may also be employed, exploiting fixation duration. On the other hand, because of the discontinuous character of saccades, Excitation is a less viable strategy.

2) EYE MOVEMENTS

Although strictly related to the previous one, this channel refers to angular displacements of the eyeballs relative to the frontal direction, with no reference to an absolute pointing direction. As such, it does not necessarily require the presence of a screen (or any fixed reference object) and has therefore the advantage of leaving freedom of movement to the subject's head and body. Regarding the properties of channel parameters, the same considerations made for Gaze pointing apply here as well.

Although vision-based eye-tracking techniques may be used, the most common sensing technology for eye movement detection is electrooculography. One specific example on the use of this channel for interacting with a HeaDMI is discussed in Sec. IV-A1. We were not able to recover additional examples in the field of assistive interfaces: in this context eye movements are always targeted at fixating points of interest on a screen, and thus fall within the gaze pointing channel discussed previously.

3) BLINKING

This physical channel refers to the vertical movement of the eyelids, which is usually impulsive (a blink). The literature makes a distinction among *voluntary* (in response to an identifiable self-initiated or external stimulus), *reflexive* (in response to a potential threat to the organism) or *spontaneous* (dependent on the psychophysical state of the individual) blinks [43]. Winks (movements of a single eyelid) have different characteristics from blinks. In particular, the ability to selectively close a single eyelid seems to be linked to personal abilities, and may be compromised by some forms of motor impairments.

Numerous blink detection algorithms based on image analysis have been developed [44], [45]. In addition eye blinks can be detected through electrooculography, as well as EEG headsets, with extensive uses for assistive technologies [45], [46]. Involuntary blinks pose problems to the interaction design, and require the ability to recognize voluntary ones. Blink duration has been used as a salient feature, where the duration of a spontaneous blink is about $300 - 350$ ms [43]. In the field of musical interfaces, this channel is still minimally explored.

The main usable parameter is the boolean (open-closed) vertical displacement of eyelids. Since an event lasts fractions of a second, this makes it suitable for the Trigger, Toggle, and Counter mapping strategies. Using the Hold strategy (associated to blink duration) has the undesirable side effect of occluding the visual channel for a prolonged time and compromising the use of other channels (e.g., gaze pointing).

4) EYEBROW MOVEMENTS

Simultaneous upward and downward movements of both eyebrows are the most straightforward to achieve, although independent movements of one eyebrow can be effectively performed by some subjects. An additional movement is squeezing, which reduces the horizontal distance between eyebrows. Spontaneous movements are present, and are known to be strongly correlated with emotional states, as well as vocal activity (particularly with prosody) [47]. Consequently, using this channel for interaction poses non trivial problems, although some studies suggest that deliberate eyebrow movements may be characterized and recognized from spontaneous ones [48]. The same studies also suggest that this channel has low resolution, and is prone to fatigue.

Correspondingly, studies in the context of accessible interfaces are almost invariably based on boolean detection of low-high eyebrow movements, e.g. to trigger mouse clicks [49]. Movements can be sensed through cameras, but also electromyography [50], and even inertial sensors such as accelerometers attached to the skin.

Similarly to eyelid movements, this channel is still minimally explored in musical interfaces (one specific example regarding HeaDMIs is discussed in Sec. IV-A14). Suitable mapping strategies are also similar and amount to Trigger, Toggle, and Counter. The Hold strategy (associated to prolonged displacements) may also be considered.

C. MOUTH

1) VOICE

The human voice is produced by a complex mechanical and acoustic system, based on the combined action of larynx (with the vocal folds), and the vocal and nasal tracts [51]. The vocal folds act as a sound source which can oscillate to produce pitched sounds, or can stay open letting the airflow through. The vocal tract acts as a filter, whose spectral characteristics are determined by the tract shape and controlled by various articulators, notably the tongue and the lips. The resonances (formants) of the vocal tract are particularly relevant to characterize vocal sounds (e.g., different vowels).

The human voice, specifically singing, is in itself a very expressive and versatile acoustic musical instrument. The use of non-verbal voice input for interactive control has been widely explored [52], [53]. In the context of musical interactions, the versatility of voice may suggest a natural mapping between singing parameters (pitch, intensity, formants, etc.) and analogous musical parameters of the sound production unit. All these parameters are suitable for Continuous range mapping strategies, and can be estimated straightforwardly through a microphone and a palette of well established signal processing techniques. Various approaches have been proposed for general purpose instruments [54]–[56].

On a different level, speech input is a well established interaction modality [57]. Speech-based control is less used in the context of DMIs, but may be suitable for Trigger mapping strategies, possibly mapped to musical parameters at the process level.

2) WHISTLING

Acoustically, a whistle is produced by the action of a Helmholtz resonator consisting of the oral cavity bounded by two orifices [58]. Anatomically, the shape of the resonant cavity responsible for the modulation of the sound is mainly given by tongue movements, rather than by jaw posture, as well as by the presence of ‘‘lateral chambers’’ inside the mouth, during the emission of notes at high frequencies [59].

Although this channel is much less explored than Voice, similar considerations may be made regarding available parameters (pitch, intensity), their related mapping strategies, and their estimation. In addition, anatomical considerations suggest that tongue position may also be used to infer pitch.

‘‘Whistling user interfaces’’ have been proposed [60]. Musical applications are mostly focused on the use of whistling as an input to query-by-humming music retrieval systems [61]. Shen and Lee proposed a whistle-to-music composition system [62], but to our knowledge this channel has not been used in DMIs.

3) BREATH

Breath is a primary interaction channel in many acoustic aerophone instruments. As a consequence it has been widely used also in DMIs, and several commercial interfaces incorporate a breath sensor. The main associated parameters (pressure and airflow) can allow for highly expressive control, as demonstrated by the variety of subtle sound nuances obtainable in acoustic instruments, and can be used in DMIs through Continuous range mapping strategies. The most typical mapping, mutuated from acoustic instruments, is between pressure and Intensity and Note on/off parameters. However different mappings may be explored.

Breath has also been considered in the context of accessibility, chiefly for the control of powered wheelchairs [63], but also for other devices, e.g., digital music players [64]. In this context breath is typically used to trigger changes in a state machine.

4) MOUTH AND LIP MOVEMENTS

The shape of lips and mouth can be controlled voluntarily, through the action of facial muscles and jaw. In ‘‘virtual human representation’’ applications (e.g., generation of avatars), parameters such as mouth aperture and mouth stretch/squeeze are typically estimated through vision-based approaches [65]. EMG sensors may also be used especially for stretch/squeeze associated to articulatory muscles on the cheek. This is demonstrated by EMG-based ‘‘silent speech’’ interfaces for speech disabled people [66]. Jaw movements may be also sensed through Hall effect sensors attached to the teeth.

In the context of musical interactions, vision-based approaches have been used to estimate and map this channel’s parameters into musical control [14], [67], mainly at the timbral level. As an example, the Mouthesizer [14] uses aperture to control timbral parameters which are applied as audio effects to instrumental sounds.

Some parameters (particularly height and area) have high resolution due to the fine control over jaw movements, which makes them suitable for Continuous range mapping strategies. Other parameters (particularly those associated to cheek muscles) are more suited for mapping strategies such as Carousel selections, Trigger, or Hold.

5) TONGUE

The tongue is capable of very rapid and precise movements [51]. It is customary to divide it into sections: tip, blade, front, back, and root. For simplicity, and in accordance with typical applications found in the literature, we limit our analysis to the 3D position of the tip. Additional parameters, such as pressure against teeth or palate, may be considered.

Tongue pointing devices have been proposed [68]. In the field of assistive technologies, several works have used tongue tip movements for the control of powered wheelchairs [69]–[71]. Many use a small magnet positioned on the tongue (glued or installed as a piercing), and a series of magnetic sensors (e.g., Hall effect sensors) placed on the mouth, to detect the distance of the magnet. The estimated tongue tip position is thus relative to the mouth and influenced by the position of the sensor.

Alternative sensing strategies have been proposed: Vaidyanathan *et al.* [72] used a microphone in the ear canal to detect pressure variations due to tongue movements and found that at least 4 tongue gestures could be accurately recognized; Cheng *et al.* [73] used an array of textile pressure sensors attached to the cheek and showed that 5 gestures could be accurately recognized.

In light of its high velocity and resolution, which make it suitable for Continuous range mapping strategies, the tongue is an interesting musical controller. This was proposed already in 1991 [74], however with limited success. Vogt *et al.* [75] developed a music controller based on tongue posture estimation via ultrasounds. More recently, Nam and DiSalvo [76] described an experiment in sonification of tongue movements via a Hall Effect Sensor. Involuntary movements (e.g., swallowing due to salivation) are a possible drawback.

6) TEETH

Lower (mandibular) teeth can be displaced from upper (maxillary) teeth through mandible movements, independently from mouth aperture: vertical, lateral, and – to a lesser extent – longitudinal displacements can be made.

Various studies explored this channel for hands-free interaction. Most of them share a common approach based on detection and recognition of "tooth clicks", i.e. clenching actions. Typical applications are directed at controlling a pointer, but also include other use cases such as initiating a process (e.g., a phone call), controlling an ongoing process, or responding to a notification.

Employed sensors include EMG sensing on the temporal muscles [77], [78], sensing of vibrations in the jawbone and skull through an accelerometer (typically positioned around the external ear) [79], [80], as well as acoustic detection of tooth clicks using contact microphones on the throat or ear [81], [82]. Some of these studies report accurate recognition of up to seven different clicks [77], which can be used to emulate a mouse. Lateral and longitudinal displacement may also be recognized, through Hall effect sensors. One further parameter, the pressure between lower and upper clenching teeth, may also be measured: an example is provided by "food simulators" [83], which use pressure sensors in between the dental arches.

Such studies suggest that this channel may be used also for musical interactions, using Trigger, Toggle, Continuous range mapping strategies. Pressure may be used for Continuous range mapping strategies. It can be expected that all the parameters have high stability and low fatigue.

D. HEAD

1) HEAD MOVEMENTS

Active head movements, especially along the three rotational degrees of freedom (yaw, pitch, roll), are actively used in several everyday interactions: in conjunction with the vestibular and visual systems for postural balance [84], as a support to vision and audition in localization and target reaching tasks [85], [86], as a mean to convey paralinguistic information in synchrony with speech utterances [87]. The related kinematics have been extensively studied [88].

Head tracking is used ubiquitously in HCI applications, including assistive technologies (as an example, powered wheelchairs operated by head gestures are common [89]). The recent rise of virtual reality technologies include head-mounted displays with integrated head-trackers. The term "virtual reality musical instruments" (VRMIs) is now used to refer to DMIs that include a simulated visual component delivered via a head-mounted display or other forms of immersive visualization [90]. However in this case head tracking is used to provide convincing immersion in the virtual environment, rather than to control musical parameters. On the other hand, some studies have explored the use of head gestures for musical control at the timbral and note levels [91], [92].

Having high resolution, accuracy, and velocity, head movements can be used with a variety of mapping strategies, including Continuous range, Hold, and Switch. Head motion produces relatively high levels of kinetic energy. It is therefore well suited for Excitation mapping strategies, to control e.g. sound intensity. Additional natural mappings may be associated to timbral parameters such as vibrato. However, the issue of fatigue associated to prolonged movements would need further investigation.

2) NECK

This channel is associated to articulation of neck muscles, which can be detected through EMG. Muscular activations leading to changes in head orientation pertain to the previous channel and have already been discussed. However, isometric contractions and relaxations of neck muscles can also be produced, with no associated head movements, and these can also be detected by EMG.

Examples of studies employing this channel for interaction are scarce. One such example is provided in the work by Hands and Stepp [93], who experimented with the use of EMG sensors on the anterior neck and on the submental surface to control the vertical displacement of a pointer in a target reaching task. Specifically, participants were instructed to produce and maintain static EMG activations at different target levels to move the icon, for various time intervals.

The suitability of this channel for musical interactions remains to be explored.

E. BRAIN

1) MENTAL STATES

This channel is related to covert aspects of user state, including latent cognitive processes (arousal, workload, etc.) and "cognitive events" (perception of errors, bluffing, surprise, etc.). These can be seen as a secondary communication channel for HCI, that enriches the interaction through implicit user information [94]. Applications include interface evaluation, adaptive systems, and neuro-feedback.

Mental states can be recognized from the EEG signal [95, Ch.7]. Typical parameters are derived from power spectrum analysis, which divides the EEG signal into frequency bands ("rhythms") and uses the power density distribution across bands as a feature set for subsequent classification. A simple related parameter is the spectral centroid. Temporal features are also used, such as Hjorth parameters (activity, mobility, and complexity) [96], which are estimated on successive windows (epochs) from the time-domain signal and its derivatives.

The detection of affective states is particularly relevant for musical applications. "Affective BCIs" [97], [98] are often based on the so-called "valence-arousal" 2D space [99] to define emotional classes of interest. The first axis defines a dimension related to emotion positivity/pleasantness, while the second one defines the degree of engagement/excitement.

Estimated cognitive workload has been used e.g. for intelligent music tutoring systems [100] and automatic accompaniment [101]. Estimated affective states have been used for automatic generation of music, for composition [102], in computer games [103], or for modulating the affective user state [104]. Direct EEG sonification has also been explored as a way of representing mental states using auditory output [105], for monitoring, diagnostics, neuro-feedback, and communication.

The main quality of this channel is that it increases the information flow without requiring conscious effort. Therefore it has low associated fatigue, and no training is required. The associated latency is significant, although variable: as an example, the duration of the epochs used to perform emotion assessment can vary from 0.5 s to 5 minutes [98]. As such, this channel is especially suited to map onto musical parameters at the process level.

2) Attention

This channel comprises brain signals evoked by external stimuli [39], [40], particularly event related potentials (ERPs). As such, it depends on attentional capacity and sensory information to be intact.

One relevant example is the P300 potential, detected in the parietal cortex ∼300 ms after the occurrence of a significant stimulus interspersed with frequent or routine stimuli. Many BCIs based on P300 use a matrix-like visual interface, operated through an "oddball paradigm" [106], in which rows and columns flash randomly: if the user focuses on a specific matrix element, a P300 peak will be produced when it flashes. The "P300 speller" is perhaps the best known embodiment of this paradigm, and allows to select letters from a matrix [107].

Steady-state visual evoked potentials (SSVEPs) are also widely used, and can be measured from the EEG at the occipital cortex during periodic presentation of visual stimuli [108], with a latency of ∼100 ms. Error-related negative potentials (ERNPs) instead occur $200 - 250$ ms after the detection of an erroneous response in a continuous stimulus-response sequence, and can be used e.g. to identify cursor movements outside a defined visual field or to detect an error in a sequence of target stimuli [109].

This channel has been used in musical interfaces. Grierson [110] proposed a "P300 composer" which allows to select and write notes using an interface based on the oddball paradigm and P300 evoked potentials. Pinegger *et al.* [111] and Chew *et al.* [112] used the P300 to select notes from a matrix. Miranda and coworkers contributed pioneering works on "Brain-Computer Musical Interfaces" (BCMIs) [95], and experimented especially with the use of SSVEPs [113], [114].

ERPs have latencies in the order of hundreds of milliseconds. Moreover their reliable detection requires many repetitions of the stimuli. This makes this channel unsuitable for triggering musical events in real time, but leaves space for process-level control. Since most working applications allow for the detection of a limited number of options, simple mapping strategies (Trigger, Toggle, Discrete range) may be used.

3) IMAGERY

This channel relates to the active performance of cognitive tasks associated to various types of mental imagery, including geometric shapes, familiar faces, tunes, word associations, calculations, and motor imagery [115]. The latter is the most common and amounts to imagining self movements, which activates primary sensorimotor areas: as an example, imagined movements of left hand cause event-related desynchronization (ERD) and synchronisation (ERS) in the right and left motor cortex, respectively [40]. These can be detected in specific bands.

A related mechanism is provided by slow cortical potentials (SCPs), which measure cortical EEG polarization related to preparation (e.g., readiness/planning to move) or decreased activation. It has been shown that a subject can learn to actively control SCPs by means of various mental tasks, and thus to use them for control [40].

Applications range from prosthetic limbs [116] to the control of quadricopter drones [117]. The most famous interface based on SCPs is probably the "Thought Translation Device" [118], which allows 1D displacement of a cursor. Some musical interfaces make use of this channel. Pham *et al.* [119] used SCPs to control pre-set pitch sequences, although these were merely intended as auditory feedback for SCP training rather than for music generation. Vamvakousis and Ramirez [120] used ERD/ERS parameters for a musical application that lets users move the pitch of a tone up and down in a musical scale. A similar approach is followed by the "encephalophone" [121], where users can generate different notes of a C major musical scale. The "Brain dreams Music" project [122] uses instead music imagery, specifically four chords, which are detected and played back.

This channel requires moderate to extensive training, both for subjects to learn and for the systems to gather sufficient data for classification. It also requires significant cognitive resources, leading to higher fatigue. Velocity is low: as an example, classification of imagined movements may require seconds. Simple mapping strategies (Trigger, Toggle, Discrete range) may be used.

IV. SURVEY OF EXISTING HeaDMIs

Building on the framework developed so far, here we review a list of HeaDMIs previously proposed either in the scientific literature or as commercial products. The literature search was performed on Google Scholar while the search for commercial products was performed on Google. Both were based on free text terms related to DMIs, namely ''digital musical instrument'', ''musical (or music) interface'', which were coupled (logical AND) to additional terms related to one of the physical channels discussed in Sec. III. Additionally, we analyzed the lists of references from previous reviews on ADMIs [8], [9].

The retrieved items were selected based on four inclusion criteria. Specifically, they had to: (i) present a skill-based DMI, according to the definition of Sec. II (thus, musical interfaces related to offline composition, sequencing, playback, etc. were not considered); (ii) make use of one or more physical channels among those discussed in Sec. III, with no additional ones (thus, neither augmented instruments nor DMIs requiring the use of upper/lower limbs were considered); (iii) describe a concrete – albeit prototypical – working implementation (thus, theoretical studies and reviews were not considered); (iv) make explicit reference to users with some form of motor impairment, among the target user groups.

As a result, a relatively small number of HeaDMIs was retrieved: a summary is provided in Table 4, while Sec. IV-A presents a structured analysis for each of them, considering the employed physical interaction channels and mappings, as well as sensors and interfaces. The degree and type of instrument evaluation (if any) are also mentioned. Based on this analysis, Sec. IV-B provides a comparative discussion of the surveyed instruments.

A. ANALYSIS

Instruments are presented in chronological order with respect to the first publication or product release.

1) BioMuse [123], [124]

Originally defined as a ''biocontroller'', this was a pioneering project (1990) and is included in this survey also for historical reasons. The system underwent several implementations, all having at their core a HW/SW developed specifically to collect EEG, EOG, and EMG signals, extract a set of relevant features, and map those to MIDI signals. Currently, a musical ensemble called ''The BioMuse Trio'' performs using BioMuse, a violin, and a laptop.

The first implementation used 3 channels: Mental states (Brain), Eye movements (Eyes), and Voice (Mouth, optional). The employed sensors were two EEG electrodes (occipital lobe and frontal area), two EOG electrodes, and a microphone. Bands with disposable snap electrodes were used. In addition the instrument included a variable number of EMGs. Although Knapp [123] discusses an example where EMGs are around biceps and forearms, the system is agnostic with respect to EMG positioning.

Possible mappings are exemplified in the first implementation. There, two EMG signals were mapped to the intensity and pitch of a synthesizer, EOG signals to stereo panning, and the Alpha component of the EEG signal to instrumental sound (from violin to glockenspiel). No extrinsic feedback is provided. The first implementation included a GUI which could be used to modify sensors thresholds and channel to MIDI mapping.

2) TONGUE-CONTROLLED ELECTRO-MUSICAL INSTRUMENT [125]

This DMI uses exclusively the Tongue (Mouth) physical channel. Tongue position is detected by a PET board extruding from the mouth, with 5 buttons (switches) placed on the palate. By virtue of the mechanical contact of tongue with buttons, tactile and proprioceptive intrinsic feedback is provided to the performer. Button activations are mapped into MIDI events.

The instrument employs a Discrete range mapping strategy with a small number of values, as well as a Counter mapping strategy. Four buttons (''scale control switches'') are arranged in a + shape and are mapped to four different chords. The fifth button (the ''chord shift switch'') is placed below them and changes the pattern of assigned chords along a circular set of available patterns, depending on the tune to be performed. This allows to explore a reasonably wide palette of chords.

The instrument was evaluated with three able-bodied subjects, mainly for the purpose of assessing rhythmic capabilities of the tongue: two subjects with previous musical experience were able to maintain a sufficiently correct rate of button depressions at three different BPMs (60, 120, 180). The song ''Twinkle, twinkle, little star'' was performed by a subject with no musical experience at 65 BPM.

3) HI NOTE [126], [127]

The instrument name refers to the most recent iteration of a commercial HeaDMI developed over several years. Previous iterations included the Headspace and the Typhoon (the year reported in Table 4 refers to the first reported release of the Headspace).

The employed channels are Head movements (Head) and Breath (Mouth). The former is mapped to pitch, discretized as a set of notes visible on the screen (Discrete range). Regarding the latter, audio-video documentation shows that the pressure parameter of Breath can be alternatively mapped to Note on/off (Toggle) or to the rate of Note on events (Continuous range), with higher pressure values producing faster ''ribattuto'' effects. The choice among these mappings is left to the performer. In a sub-section of the graphical interface, the same channels are mapped to additional control at the process level (octave transposition). Head movements are tracked using a 9-axis sensor equipped with a 3DoF Accelerometer, a 3DoF Gyroscope, and a 3DoF

TABLE 4. List of surveyed HeaDMIs. Refer to Tables 3, 1, 2 for sensors, musical parameters, and physical channels, respectively.

Instrument	References	Year	Sensors	Mapping (channel → musical param.)
BioMuse	[123], [124]	1990	eeg, eog, emg	ment→isel ; eymv→pann(stereo)
Tongue-Controlled EMI	[125]	2004	btn	tong→pitc/hcha,note
Hi Note	[126], [127]	2005	head, brea	head→pitc ; brea→note
Magic Flute	[128]	2007	gyro, brea	head→pitc ; brea→note,inte
Jamboxx	[129]	2009	?, brea	head→pitc ; brea→note,inte
Lumiselo	[130]	2010	eyet, brea	gaze→pitc ; brea→note,inte
SSVEP BCMI	[131], [132]	2011	eeg	atte→pitc
EyeHarp	[133]–[135]	2012	eyet, eeg	gaze→ pitc,note,inte,vibr,hcha ; ment→temp,msel
Eye play the Piano	[136]	2014	eyet	gaze→pitc,note,hcha
P300 Harmonies	[137]	2014	eeg	atte→hcha
Imitone	[138]	2014	micr	[voic,whis]→ pitc,note,inte,brig,glid,vibr
Clarion	[139]	2015	eyet	[head,gaze]→note,pitch,vibr,brig
EyeJam	[140]	2015	eyet	gaze→pitc,note
Eye conductor	[141]	2016	eyet, cam	gaze→pitc,note ; eybr→tran(octave) ; mout→brig
Netytar	[142]	2018	eyet, brea	gaze→pitc ; brea→note,inte

Magnetometer, which is claimed to provide accurate and high-resolution tracking. Mouth pressure is detected by a breath sensor.

No formal evaluation studies were retrieved. However, the Hi Note is used in public performances and notably it has featured in the 2012 Paralympic Games closing ceremony, with the British Paraorchestra.

4) MAGIC FLUTE [128]

This commercial DMI uses Head movements (Head) and Breath (Mouth) as physical channels. The instrument physically consists of a mouthpiece rigidly connected to the main instrument body. The body swivels on a standard camera/microphone mount, and is hinged on the lateral axis. Consequently it follows head rotations (pitch) of the performer holding the mouthpiece. Thus, head position is inferred through mechanical contact with the instrument.

The vertical rotation of the instrument (head rotation along pitch angle) is detected by an embedded gyroscope and is mapped to musical pitch, discretized along the musical scale (Discrete range). Mouth pressure, detected by a breath sensor, controls Note on/off events (Toggle) and intensity (Continuous range). An additional control module allows the performer to choose from a set of predefined instrumental timbres.

The instrument provides no extrinsic feedback to the performer. No formal evaluation studies were retrieved. However, the Magic Flute is notably used in the musical activities of the Dutch foundation "My Breath My Music", devoted to people with motor impairments.

5) JAMBOXX [129]

This is a commercial DMI which works very similarly to an acoustic mouth harmonica. It shares several features with the Magic Flute and uses Head movements (Head) and Breath (Mouth) as physical channels. It consists of a horizontal body that swivels on a standard camera/microphone mount and is equipped with a movable mouthpiece. This can slide along the lateral axis of the body and consequently it follows head rotations (yaw) of the performer holding the mouthpiece. Additionally the instrument can also rotate vertically along the lateral axis.

Yaw and pitch head rotations are detected through the lateral displacement of the mouthpiece and the vertical rotation of the instrument (details of the sensors are not provided). The former parameter is mapped to pitch discretized along the musical scale (Discrete range), like in a mouth harmonica, while the latter is used to add pitch glides ("pitch bend" effects, Continuous range). Mouth pressure, detected by a breath sensor, controls Note on/off events (Toggle) and intensity (Continuous range).

Tactile feedback is provided by means of "bumps" (crests and troughs) on the instrument face which, in analogy to frets on a guitar, provide information about horizontal position. Vogels [143] evaluated the usability of the instrument, with five able-bodied subjects.

6) LUMISELO [130]

Presented by the authors as an "electronic wind instrument", it makes use of Gaze pointing (Eyes) and Breath (Mouth) as physical channels. It consists of head-mounted goggles equipped with a custom-made eye-tracker (infrared LED and camera), as well as a breath sensor connected to the performer's mouth by a rubber tube. The authors emphasize that this design allows free head movements to have no effect on the location of the performers pupil with respect to the visor (unlike gaze pointing interfaces based on external monitors).

A 12 × 3 grid of LEDs on the visor represents three octaves of a chromatic scale and the corresponding 2D coordinates of the gaze point are mapped to pitch (Discrete range). Mouth pressure controls Note on/off events (Toggle) and intensity (Continuous range).

The LED pointed by the performer's gaze changes color on the visor. Additionally, the pressure detected by the breath sensor controls the brightness of the same LED. This provides extrinsic visual feedback, which is claimed to improve performance and engagement (although no formal evaluation is presented).

7) SSVEP BCMI [131], [132]

One of the many BCMIs developed by Miranda and coworkers, this was specifically designed for performers with severe motor impairments, and was tested with a patient with Locked-in Syndrome. The instrument allows for real-time generation of melodic lines using Attention (Brain) as the only channel.

The first implementation employed a visual interface with four icons, and SSVEPs (sensed through EEG) were used to detect the selected icon and the intensity with which that icon was attended. Visual feedback was provided by changing the size of the icon in relation to the magnitude of the SSVEP signal. Various subsequent implementations were released [132].

The employed mapping is a distinguishing feature of this instrument. Within a given frequency band (associated to the flashing rate of one of the icons), the signal magnitude is quantized to five values, each corresponding to a note from a predefined 5-note pattern. In this way the SSVEP parameter is mapped to pitch (Discrete range), allowing the performer to generate melodies.

The authors report [131] on their trial with the Locked-in Syndrome patient was successful in terms of usability and engagement. In particular, response times between attending a target and the corresponding musical event were approximately $1 - 2$ s.

8) EyeHarp [133]–[135]

This instrument is based exclusively on Gaze pointing (Eyes), but realizes a complex interaction by means of a two-layer interface: one layer manages the performance of melodic lines, while the second one allows to compose short accompanying patterns that can be played in loop.

The "melody layer" employs a pie-shaped visual interface, with slices associated to pitches (Discrete range), and with an inactive area in the center. This layout (together with the introduction of a dwell-time to recognize fixations) alleviates the Midas touch problem. The radial position inside a slice is mapped to intensity and vibrato (Continuous range): an example of a one-to-many mapping.

Visual feedback is minimalistic yet informative, and amounts to the appearing of one or more focus points in the selection area of a selected pitch. The instrument has been extensively evaluated from the perspective of both the performer and the audience [134].

A fork of the project led to a "Brain-Gaze controlled" musical interface augmented with an additional channel, namely Mental states (Brain), where emotional states control the "step sequencer layer". Valence is mapped to three different chord sequences (Discrete range), while arousal is mapped to the relative intensities of four predefined arpeggios (Continuous range). Additional visual feedback is provided in the form of varying colors (associated to valence) and brightness (associated to arousal).

9) EYE PLAY THE PIANO [136]

This DMI was developed collaboratively between a commercial manufacturer of VR headsets and the University of Tsukuba. Despite the lack of accompanying publications, the main characteristics can be deduced from the project web site and audio-video materials.

The distinguishing feature is the use of an actuated acoustic piano, which can then be played without hands. Gaze pointing and Blinking (Eyes) are employed as physical channels, by means of an eye tracker integrated into a HMD. This allows free head movements of the performer.

Two different mappings are implemented for gaze pointing. The "monotone" and "chord" modes map to piano notes pitches, and to a set of chords, respectively: in both cases, a Discrete range mapping strategy is adopted. Blinking is mapped to Note on events (Trigger strategy): a pointed pitch or chord is triggered by a blink, and a corresponding Note off event is produced when a subsequent blink triggers a different pitch or chord.

Visual feedback is provided through the stereoscopic HMD display. Colored keys are placed on a virtual surface and mapped to single notes or chords. Selection of a key is signalled by visual effects. The number and type of keys displayed on the surface can be customized.

No formal evaluation appears to have been conducted. However the instrument has been used in public performances by young performers with motor impairments. At the time of writing the project is supported through charity fundraising.

10) P-300 HARMONIES [137]

This DMI focuses on real-time generation and modification of arpeggios. The employed physical channel is Attention (Brain), and P300 evoked potentials are detected through a low-cost, 14-channel commercial EEG device.

Arpeggios in the instrument consist of loops of 6 notes, played in random order to trigger the P300 response, with predefined timing and duration. A visual interface shows a 2×6 matrix, whose columns are "switches", i.e. vertical pairs of notes which correspond to "up" and "down" switch positions. Each switch flashes when the corresponding note is played. The visual interfaces then serves both as extrinsic feedback and as a provider of stimuli for P300 responses.

By focusing on a specific switch, the performer can toggle between the up and down positions, thus modifying one note of the arpeggio. The selected note of the switch is highlighted with a different color on the visual interface. The P300 channel parameter is therefore mapped onto pitch, using a Discrete range strategy with a limited number of options.

Preliminary evaluation of the instrument was carried out with 4 subjects, in terms of accuracy in performing the toggle action on each of the 6 switches.

11) IMITONE [138]

This commercial DMI is a software app, which makes use of Voice (Mouth) as the only physical channel. Whistling (Mouth) is also mentioned as a possible alternative channel. The only required sensor is a microphone.

The main advertised feature is the use of an efficient and accurate pitch tracking algorithm, which allows for low-latency conversion of voice parameters into MIDI events and consequently for voice-based real-time control of virtual instruments.

Several parameters are used for Note-level control, including Pitch, Voice activity, and Intensity, which are mapped to corresponding MIDI messages. Pitch in particular allows for either Discrete range (pitch quantization on various scales) or Continuous range strategies (glissando). In addition the instrument allows for control at the Timbral level, by detecting vibrato and glides.

A graphical interface provides extrinsic visual feedback in particular with regard to pitch, displayed through either a ‘‘piano roll’’ visualization or a continuous plot.

One advantage of using the Voice channel and intuitive mappings is that the instrument has minimal requirements in terms of expertise and training. However no formal evaluation studies were retrieved.

12) CLARION [139]

This is a DMI developed through a long-term charity program, with a strong emphasis on participatory design, adaptability to individual needs, and exploitation of off-the-shelf assistive technologies used by persons with disabilities in their everyday lives.

It can use various alternative physical channels, including Gaze pointing and Head movements (both detected via a commercial eye tracker), but can also be played with one’s fingers or feet. The chosen channel is mapped to Note on/off and to Pitch, via a graphical interface that represents notes on a screen (Discrete range strategy). Additionally, timbral parameters can be controlled by the rate of movement of the physical channel (Excitation strategy) and by the position within the area representing a single note (Continuous range strategy): an example of a one-to-many mapping.

Notes can be arranged into several different layouts and assigned different shapes, sizes, and colours. This, together with the possibility of choosing among a palette of physical channels, allows for high adaptability. Visual effects in the graphical interface provide extrinsic feedback to the performer. Although no formal evaluation studies were retrieved, the instrument is actively used by a large number of performers through the Open Orchestras initiative.

13) EyeJam [140]

This DMI allows for the generation of melodic lines using Gaze pointing (Eyes) as the only physical channel. It exploits an interesting ‘‘context switching’’ interaction paradigm, proposed by the same authors, which addresses the Midas touch problem by associating focus and selection to different eye movements.

Specifically, the visual interface is made of two identical keyboards placed at the top and bottom of the screen (the two ‘‘contexts’’) and separated by a narrow empty area (a ‘‘bridge’’): in order to select a key, the gaze trajectory must cross the bridge and switch context through a saccadic movement, whereas fixations on different keys within the same context do not trigger any event.

The 2D gaze point position on the screen is mapped to pitch of the keyboard keys, using a Discrete range mapping strategy. A low-cost commercial eye tracker is used, whose limitations in terms of accuracy and calibration are dealt with by providing a limited number of large keys: 9, arranged along a major scale. Keys are color coded according to their position on the scale. The visual interface can also provide additional feedback to help a user follow and learn a predefined melody.

The system was evaluated with 6 able-bodied subjects, who compared the proposed interface with one where selection is based on dwell-time. Experimental results suggest that the context switching paradigm allows for improved accuracy in rhythmic tasks.

14) EYE CONDUCTOR [141]

This DMI is notable for its use of several physical channels, including Gaze pointing (Eyes), Eyebrow movements (Eyes), and Mouth-lip movements (Mouth), employed for control at the process, note, and timbral level. Low-cost sensors (a commercial eye-tracker and a webcam) are used.

Gaze pointing is mapped to Pitch (Discrete range) and Note on/off, using a pie-shaped interface in which eight sectors represent an octave, and a central ‘‘silent area’’ corresponds to Note off events (this layout shares some similarities with EyeHarp). Although not specified in the documentation, selection appears to be based on dwell-time. Eyebrow movements are mapped to transposition using a Toggle strategy: raising/lowering eyebrows transposes all pitches by one octave up/down. Mouth aperture is mapped to timbral brightness (other possibilities are envisioned, such as controlling reverb or delay-based effects). Alternative layouts are also proposed, which include control over chords and drum sequencing.

The graphical interface provides extrinsic visual feedback to the performer. Circle sectors are colored upon selection of the corresponding notes, and a stylized animated silhouette of the performer’s face is rendered in the background.

There are no accounts of formal evaluation. However the accompanying audio-video materials show that several preliminary user tests were conducted with people with various motor impairments, since the early design stages. Notably, the documentation mentions an adaptation step that allows for adjusting thresholds for facial gestures to fit abilities of different users.

15) NETYTAR [142]

This monophonic DMI uses Gaze pointing (Eyes) and Breath (Mouth) as physical channels. The former is mapped to pitch (Discrete range), by means of a 2D graphical interface. Mouth pressure, detected by a breath sensor, controls Note on/off events (Toggle) and intensity (Continuous range). The employed physical channels and mappings share several similarities with Lumiselo.

The main distinguishing feature is the layout of the 2D graphical interface, where discrete pitches are represented as nodes of a 2D grid and each is connected by arcs to 6 adjacent nodes: different directions correspond to different musical intervals. This layout is qualitatively resemblant of strings and frets on a guitar, where notes can be reached by moving along frets and across strings. An "isomorphic" property holds: a trajectory on the grid associated to a melodic line is independent on pitch transpositions. This design is claimed to alleviate the Midas touch problem by creating direct connections between notes for the most common musical intervals.

The interface provides extrinsic visual feedback: nodes are colored according to their pitch, and the grid scrolls on the screen to keep the gaze point always in proximity of the center (taking advantage of eye's smooth pursuits). This makes the grid and instrumental range potentially infinite.

The instrument was evaluated with eight able-bodied subjects, who performed simple melodies using both the Netytar and the EyeHarp. Quantitative evaluations (average timing errors, number of wrong notes) suggest comparable playabilities of the two instruments. Qualitative evaluations (usability questionnaire) were also presented.

B. DISCUSSION

Most of the HeaDMIs surveyed in the previous section were developed in recent years, suggesting a growing interest from both the research community and the industry. It is worth noting that a significant portion of the surveyed instruments are either commercial products or non-academic projects funded through charity programs: this suggests that there is a potential for exploitation and a large population of target users. At the same time, the limited number of surveyed instruments, as well as their design choices, shows that the field is still underexplored.

In order to provide a structured and visually-oriented comparison of the instruments, we resort to the dimension space analysis approach proposed by Birnbaum *et al.* [27]. This HCI-driven approach has been used to evaluate the main design aspects of a broad range of "musical devices" (musical instruments, interactive installations, games, and so on) along the 7 axes depicted in Fig. 3. Required Expertise represents the level of practice of the performer; Musical Control specifies the level of control exerted by the performer (see Sec. II-B); Degrees of Freedom indicates the number of independent available channel parameters; Feedback Modalities indicates the degree to which extrinsic feedback is provided; Inter-actors represents the number of people involved in the musical interaction; Distribution in Space represents the total physical area in which the interaction takes place; Role of Sound represents the category of sound role.

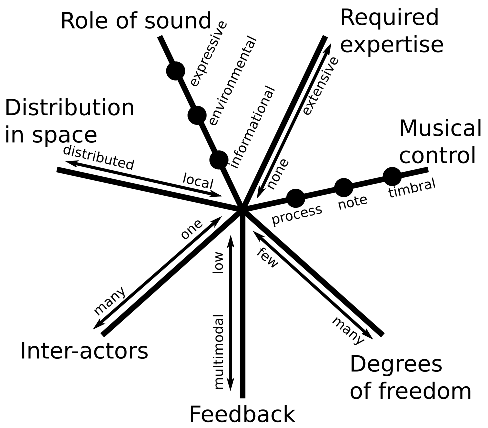

FIGURE 3. The 7-axis dimension space proposed by Birnbaum *et al.* [27] for the analysis of musical devices.

Fig. 4 represents the surveyed instruments on the Birnbaum dimension space. Even though such representation involves qualitative and subjective evaluations, it highlights some relevant points. Values on the three left-hemiplane axes are shared by all the instruments ("expressive" role of sound, "local" distribution in space, "one" inter-actor). These axes are included here for the sake of compliance with the original formulation, which is meant to represent other types of musical devices in addition to skill-based instruments considered here. Instead, relevant differences and common trends may be observed along the four axes in the right hemiplane.

Most of the instruments have few degrees of freedom, corresponding to low numbers of employed physical channel parameters. This can be also appreciated from the data reported in Tab. 4, according to which the explored channels are Gaze pointing (7 instruments), Breath (5), Head movements (4), Mental states (2), Attention (2), Voice (1), Tongue (1), Eye movements (1), Mouth and lip movements (1), EyeBrows (1), Whistling (1).

The musical control exerted by the performer is generally confined at the note level or even at the process level (generation of chords, arpeggios, etc.). This correlates with the generally low number of degrees of freedom. The only exceptions in this respect are Imitone, Clarion, and Eye Conductor, in which control at the timbral level is achieved to some extent, also by means of a larger number of degrees of freedom.

The variety of mappings is also limited. As an example, all the instruments using gaze pointing or head movements map parameters of these channels into pitch, typically through a Discrete range strategy. Similarly, breath (pressure) is always mapped into intensity through a Continuous range strategy. Brain channels are always mapped to musical parameters

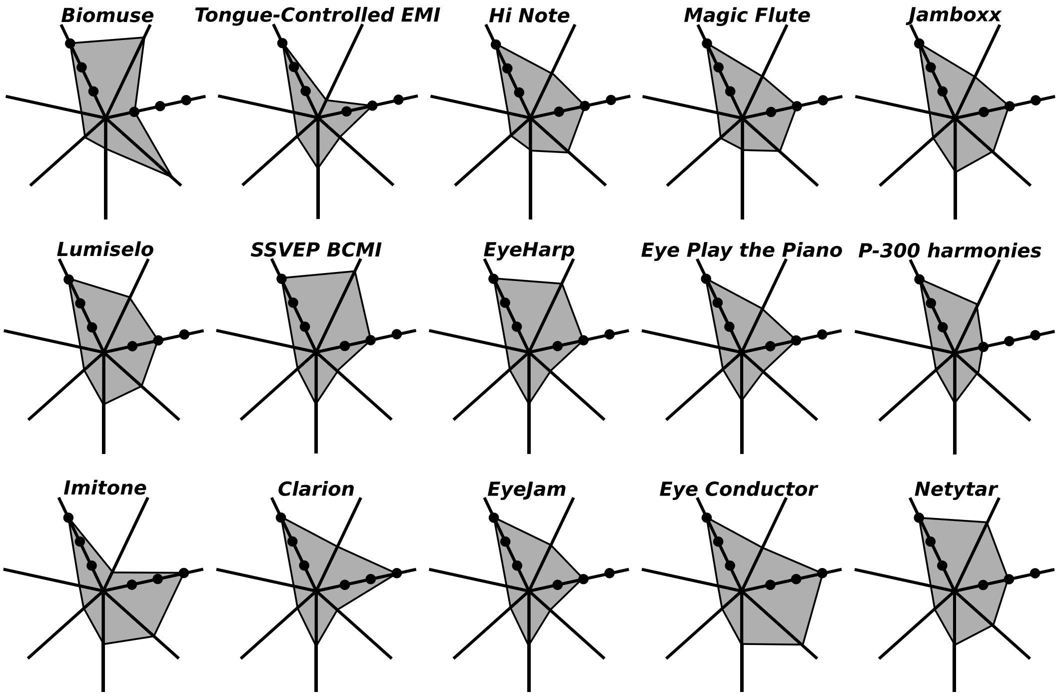

FIGURE 4. Dimension space analysis of the reviewed HeaDMIs (refer to Fig. 3 for the meaning of the axes).

at the process level (with the notable exception of SSVEP BCMI). Also, all the mappings are of the one-to-one type, in which a single channel parameter influences a single musical parameter, whereas more complex cardinalities are rarely explored. Notable exceptions are Eyeharp and Clarion, which both provide examples of one-to-many mappings.

All the instruments score low values along the Feedback axis. Apart from intrinsic and auditory feedback, additional extrinsic feedback is absent or very limited (most typically, 2D visuals on a screen). Moderately higher levels of feedback are provided by Lumiselo and Eye Play The Piano (3D visual feedback on helmet or Head-Mounted Display), while Jamboxx (tactile bumps) and Tongue-Controlled EMI include forms of tactile and proprioceptive feedback.

Several instrumental designs are lacking extensive and structured evaluation based on frameworks commonly used for DMIs (see Sec. II-E). In the absence of structured evaluation, the levels of required expertise plotted in Fig. 4 are estimated qualitatively based on our subjective judgement, and vary considerably depending on the employed parameters, mappings, and interfaces.

V. CHALLENGES AND PERSPECTIVES

The survey and the structured comparison reported in the previous section provide the ground for discussing current limitations in the design and development of HeaDMIs, presenting related open challenges, and proposing future directions of research.

A. CHANNELS, MAPPINGS, FEEDBACK

Only a small subset of potentially available physical channels is used in the surveyed instruments. Some are rarely considered, and some are completely ignored (Blinking, Teeth, Neck Tension, Imagery). A more comprehensive exploration of alternative channels is an endeavour for future work. The set of channels proposed in Sec. III could also be extended: as an example, the relatively subtle movements of nose and ears could be considered as a possible source of input control for certain individuals with motor impairments, yet no previous related studies could be retrieved at the time of writing.

In addition, physical channels and related parameters need to be characterized in terms of properties providing useful indicators for musical interactions. We proposed a set of such properties (Table 2, 9 rightmost columns), which may be reconsidered or expanded. However this issue remains largely untouched in DMI-related research, with few exceptions reported in Sec. III: whereas the importance of characterizing sensors for musical applications is well recognized [30], the same cannot be said for intrinsic characteristics of physical channels, especially unconventional ones discussed in this work.

Concerning control and mappings, our survey shows that musical control rarely extends to the timbral level, which limits the expressive possibilities of the instruments to a great extent. The predominance of one-to-one mappings is also a major limitation for instrumental expressivity. As already discussed, mappings with higher cardinalities (many-to-one, one-to-many, many-to-many) are typical of most acoustic instruments and have the potential to be more rewarding and to provide more expressive control. Finally, the generally limited (or absent) extrinsic feedback impacts negatively especially on the transparency of the interaction, for both the performer and the audience. All these aspects should be considered together in the design of future HeaDMIs.

B. ADAPTATION AND INTELLIGENCE

Most of the reviewed instruments allow for a limited degree of adaptation to different needs of various groups of users. Some include the possibility of customizing parts of the interface and musical features (e.g., range). For an instrument designed for users with different types of motor impairments, however, a key asset would be the possibility of adapting the employed channel parameters and the mappings: an example in this direction is provided by Clarion, which allows the use of gaze pointing or touch, for partially able-bodied musicians.

A recent trend amounts to using machine learning techniques in order for an instrument to learn preferred or idiosyncratic gestures of a user, and to map these to musical parameters. A notable example is the Wekinator software [144], in which various supervised machine learning approaches are used to build musical mappings through training examples. Interestingly, this software has been used in a recent project aimed at building customized musical rehabilitation devices for children with severe motor impairments [145].

Further insights can be found in the related emerging field of Smart Musical Instruments (SMIs), which can be defined as instruments equipped with embedded intelligence and able to communicate with external devices. Specifically, the five abilities of a SMI identified by Turchet [146] comprise in particular (i) context awareness, including models of the performer (needs, goals, state, etc.), (ii) reasoning, including sensor fusion approaches to define control mappings, (iii) learning, including learning from the way a performer interacts with the SMI, and (iv) adaptation and proactivity, e.g. exploiting knowledge about the performer to adapt its function or behavior. Related design principles include in particular (i) personalization, and (ii) embedded intelligence, which however maintains the performers sense of control. It is argued that the main current obstacle to the creation of SMIs is the lack of hardware and software tools able to guarantee low-latency in conjunction with all activities related to knowledge management, reasoning, and learning (feature extraction from audio and sensors, other forms of sensor signal processing, sensor fusion and machine learning, etc.). This latter remark may be extended to the fields of ADMIs and HeaDMIs, and may explain why the issue of adaptation is largely disregarded in almost all of the instruments reviewed in this work.

C. DESIGN AND EVALUATION

We have reasoned in Sec. II-E about the multifaceted nature of evaluation in the context of DMIs and HeaDMIs in particular. This is counterpointed by the lack of structured evaluation for most of the surveyed instruments. The development of evaluation frameworks specifically devised for ADMIs and HeaDMIs is certainly a challenge for future research. One further open issue is a general lack of musical pedagogies and repertoire for these instruments, which not only hinders their adoption and longevity [147], but also limits the possibility of conducting longitudinal studies targeted at long-term evaluation (learning, retention, and so on).

A related point is about the approach to instrument design. All stakeholders should be involved in the design process since the early stages, using a cyclical, participatory design approach [148] in which mock-ups and early prototypes are evaluated and redesigned based on stakeholders' feedback. This is even more needed in the case of HeaDMIs, where target users have specific and individual needs and requirements.

Principles of participatory design have only recently started to enter the mainstream of DMI research [149]. Although some of the surveyed instruments mention the involvement of one or more subjects in the design cycle, developing participatory design approaches specific to HeaDMIs is yet another challenge for future research.

VI. CONCLUSION

The contribution of this work is twofold. First, we proposed a modular and adaptable conceptual framework for the design of accessible digital musical instruments, targeted at performers with motor impairments, and based on unconventional physical interaction channels on head, mouth, eyes, brain: starting from a general structural diagram, we revisited a set of definitions and design issues related to digital musical instruments and we proposed a list of physical interaction channels available from the neck upwards, each of which was discussed in terms of potential and limitations for musical interactions. Second, we presented a systematic survey of previously developed instruments: each was analyzed in terms of design choices, physical interaction channels and related sensors, mapping strategies, performer interface and feedback. This survey provided the basis for discussing common trends and research challenges.

We have shown that the majority of the available channels is under-explored or completely unexplored. Many of these have have potential for musical interactions, therefore future works should aim at exploiting them. Similarly, there is space for exploring different mappings strategies and more complex mapping cardinalities. A particularly important challenge is the development of more expressive instruments that allow for finer control of musical parameters at the timbral level. Increased attention should be devoted to new ways to provide multimodal feedback to the performer, also considering possible losses of sensory functions that may accompany motor impairments. The ultimate goal is adaptability, which could be achieved by a completely modular framework that allows to easily and quickly create expressive tools, fully adaptable to the performer's able channels, including customizable interfaces and mappings. This calls for extensive evaluation of both instruments and physical channel properties, as well as an increased use of participatory design principles.

REFERENCES

[1] I. Cross, "The nature of music and its evolution," in *The Oxford Handbook of Music Psychology*, S. Hallam, I. Cross, and M. Taut, Eds. Oxford, U.K.: Oxford Univ. Press, 2009, pp. 3–13.

[2] H. Honing, O. Ladinig, G. P. Háden, and I. Winkler, "Is beat induction innate or learned?: probing emergent meter perception in adults and newborns using event-related brain potentials," *Ann. New York Acad. Sci.*, vol. 1169, no. 1, pp. 93–96, Jul. 2009.

[3] E. Costa-Giomi, "Effects of three years of piano instruction on Children's academic achievement, school performance and self-esteem," *Psychol. Music*, vol. 32, no. 2, pp. 139–152, Apr. 2004.

[4] K. Stensaeth, "'Musical co-creation'? Exploring health-promoting potentials on the use of musical and interactive tangibles for families with children with disabilities," *Int. J. Qualitative Stud. Health Well-Being*, vol. 8, no. 1, P. 20704, Jan. 2013.

[5] E. R. Miranda and M. M. Wanderley, *New Digital Musical Instruments: Control and Interaction beyond the Keyboard*. Middleton, WI, USA: AR Editions, 2006.

[6] A. R. Jensenius and M. J. Lyons, *A NIME Reader: Fifteen Years of New Interfaces for Musical Expression*. Cham, Switzerland: Springer, 2017.

[7] A. McPherson, F. Morreale, and J. Harrison, "Musical instruments for novices: comparing NIME, HCI and Crowdfunding approaches," in *New Directions in Music and Human-Computer Interaction*, S. Holland, T. Mudd, K. Wilkie-McKenna, A. McPherson, and M. M. Wanderley, Eds. Cham, Switzerland: Springer, 2019, ch. 12, pp. 259–267.

[8] E. Frid, "Accessible digital musical Instruments—A review of musical interfaces in inclusive music practice," *Multimodal Technol. Interact.*, vol. 3, no. 3, p. 57, Jul. 2019.

[9] J. V. Larsen, D. Overholt, and T. B. Moeslund, "The prospects of musical instruments For people with physical disabilities," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*. Brisbane, OLD, Australia: Griffith Univ., Jul. 2016, pp. 327–331.

[10] K. Graham-Knight and G. Tzanetakis, "Adaptive music technology: History and future perspectives," in *Proc. Int. Computer Music Conf. (ICMC)*, Denton, TX, USA, 2015, pp. 416–419.

[11] J. Harrison and A. McPherson, "An adapted bass guitar for one-handed playing," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, Copenhagen, Denmark, May 2017.

[12] W. Penfield and T. Rasmussen, *The Cerebral Cortex of Man; a Clinical Study of Localization of Function*. New York, NY, USA: Macmillan, 1950.

[13] S. K. Card, J. D. Mackinlay, and G. G. Robertson, "A morphological analysis of the design space of input devices," *ACM Trans. Inf. Syst. (TOIS)*, vol. 9, no. 2, pp. 99–122, Apr. 1991.

[14] M. J. Lyons, M. Haehnel, and N. Tetsutani, "Designing, playing, and performing with a vision-based mouth interface," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, Montreal, QC, Canada, 2003, pp. 116–121.

[15] A. Sears, M. Young, and J. Feng, "Physical disabilities and computing technologies: An analysis of impairments," in *The Human-Computer Interact. Handbook*, 2nd ed., A. Sears and J. A. Jacko, Eds. New Jersey, NJ, USA: Lawrence Erlbaum Associates, 2008, ch. 42, pp. 829–852.

[16] Y. Kang, H. Ding, H. Zhou, Z. Wei, L. Liu, D. Pan, and S. Feng, "Epidemiology of worldwide spinal cord injury: A literature review," *J. Neurorestoratol.*, vol. 6, pp. 1–9, Dec. 2017.

[17] W. H. Organization, "Spinal cord injury report," W. H. Org., Geneva, Switzerland, Tech. Rep., Nov. 2013. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/spinal-cord-injury

[18] A. Singh, L. Tetreault, S. Kalsi-Ryan, A. Nouri, and M. G. Fehlings, "Global prevalence and incidence of traumatic spinal cord injury," *Clin. Epidemiol.*, vol. 6, pp. 309–331, Sep. 2014.

[19] E. O. Talbott, A. M. Malek, and D. Lacomis, "The epidemiology of amyotrophic lateral sclerosis," in *Handbook of Clinical Neurology*, vol. 138. Amsterdam, The Netherlands: Elsevier, 2016, pp. 225–238.

[20] A. Chiò, G. Logroscino, B. J. Traynor, J. Collins, J. C. Simeone, L. A. Goldstein, and L. A. White, "Global epidemiology of amyotrophic lateral sclerosis: A systematic review of the published literature," *Neuroepidemiology*, vol. 41, no. 2, pp. 118–130, 2013.

[21] O. Hardiman, A. Al-Chalabi, C. Brayne, E. Beghi, L. H. van den Berg, A. Chio, S. Martin, G. Logroscino, and J. Rooney, "The changing picture of amyotrophic lateral sclerosis: Lessons from European registers," *J. Neurol., Neurosurg. Psychiatry*, vol. 88, no. 7, pp. 557–563, Jul. 2017.

[22] G. Dubus and R. Bresin, "A systematic review of mapping strategies for the sonification of physical quantities," *PLoS ONE*, vol. 8, no. 12, Dec. 2013, Art. no. e82491.

[23] J. Malloch, D. Birnbaum, E. Sinyor, and M. M. Wanderley, "Towards a new conceptual framework for digital musical instruments," in *Proc. Int. Conf. Digit. Audio Effects (DAFx)*, Montreal, QC, Canada, 2006, pp. 49–52.

[24] J. Rasmussen, "Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-13, no. 3, pp. 257–266, May 1983.

[25] P. McGlynn, "Interaction design for digital musical instruments," Ph.D. dissertation, Dept. Music, National Univ. Ireland, Maynooth, Republic of Ireland, May 2014.

[26] W. A. Schloss, "Recent advances in the coupling of the language MAX with the matthews/boie radio drum," in *Proc. Int. Comput. Music Conf. (ICMC)*, Glasgow, U.K., 1990, pp. 390–398.

[27] D. Birnbaum, R. Fiebrink, J. Malloch, and M. M. Wanderley, "Towards a dimension space for musical devices," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, Vancouver, BC, Canada, 2005, pp. 192–195.

[28] S. Fels, A. Gadd, and A. Mulder, "Mapping transparency through metaphor: Towards more expressive musical instruments," *Organised Sound*, vol. 7, no. 2, pp. 109–126, Aug. 2002.

[29] A. Hunt, M. M. Wanderley, and M. Paradis, "The importance of parameter mapping in electronic instrument design," *J. New Music Res.*, vol. 32, no. 4, pp. 429–440, Dec. 2003.

[30] M. M. Wanderley and P. Depalle, "Gestural control of sound synthesis," *Proc. IEEE*, vol. 92, no. 4, pp. 632–644, Apr. 2004.

[31] R. A. Magill and D. I. Anderson, *Motor Learning and Control: Concepts and Applications*, vol. 11, 11th ed. New York, NY, USA: McGraw-Hill, 2017.

[32] S. O'Modhrain, "A framework for the evaluation of digital musical instruments," *Comput. Music J.*, vol. 35, no. 1, pp. 28–42, Mar. 2011.

[33] O. Capra, F. Berthaut, and L. Grisoni, "Toward augmented familiarity of the audience with digital musical instruments," in *Proc. Int. Symp. Comput. Music Multidisciplinary Res. (CMMR)*, Matosinhos, Portugal, 2017, pp. 558–573.

[34] L. Turchet, T. West, and M. M. Wanderley, "Touching the audience: Musical haptic wearables for augmented and participatory live music performances," *Pers. Ubiquitous Comput.*, Mar. 2020, doi: 10.1007/s00779-020-01395-2.

[35] S. Keates, J. Clarkson, and P. Robinson, "Investigating the applicability of user models for motion-impaired users," in *Proc. 4th Int. ACM Conf. Assistive Technol. (Assets)*, 2000, pp. 129–136.

[36] P. Majaranta and A. Bulling, "Eye tracking and eye-based human–computer interaction," in *Advances in Physiological Computing*, S. H. Fairclough and K. Gillead, Eds. London, U.K.: Springer, 2014, pp. 39–65.

[37] T. S. Saponas, D. S. Tan, D. Morris, and R. Balakrishnan, "Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces," in *Proc. 26th Annu. CHI Conf. Hum. Factors Comput. Syst. (CHI)*, 2008, pp. 515–524.

[38] W. Hess, "Head-tracking techniques for virtual acoustics applications," in *Proc. Audio Eng. Soc. Conv. (AES)*, 2012, p. 8782.

[39] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain–computer interfaces for communication and rehabilitation," *Nature Rev. Neurol.*, vol. 12, no. 9, pp. 513–525, 2016.

[40] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophys.*, vol. 113, no. 6, pp. 767–791, 2002.

[41] C. Guger, B. Z. Allison, and N. Mrachacz-Kersting, "Brain-computer interface research: A state-of-the-art summary 7," in *Brain-Computer Interface Research*. Cham, Switzerland: Springer, 2019, pp. 1–9.

[42] A. J. Hornof, "The Prospects For Eye-Controlled Musical Performance," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*. London, U.K.: Goldsmiths Univ., Jul. 2014.

[43] F. VanderWerf, P. Brassinga, D. Reits, M. Aramideh, and B. Ongerboer de Visser, "Eyelid movements: Behavioral studies of blinking in humans under different stimulus conditions," *J. Neurophysiol.*, vol. 89, no. 5, pp. 2784–2796, May 2003.

[44] K. Arai and R. Mardiyanto, "Comparative study on blink detection and gaze estimation methods for HCI, in particular, Gabor filter utilized blink detection method," in *Proc. 8th Int. Conf. Inf. Technol., New Generat.*, Las Vegas, NV, USA, Apr. 2011, pp. 441–446.

[45] K. S. Park and K. T. Lee, "Eye-controlled human/computer interface using the line-of-sight and the intentional blink," *Comput. Ind. Eng.*, vol. 30, no. 3, pp. 463–473, Jul. 1996.

[46] A. Krolak and P. Strumillo, "Vision-based eye blink monitoring system for human-computer interfacing," in *Proc. Conf. Human Syst. Interact.*, May 2008, pp. 994–998.

[47] I. Guaïtella, S. Santi, B. Lagrue, and C. Cavé, "Are eyebrow movements linked to voice variations and turn-taking in dialogue? An experimental investigation," *Lang. Speech*, vol. 52, nos. 2–3, pp. 207–222, Jun. 2009.

[48] K. L. Schmidt, S. Bhattacharya, and R. Denlinger, "Comparison of deliberate and spontaneous facial movement in smiles and eyebrow raises," *J. Nonverbal Behav.*, vol. 33, no. 1, pp. 35–45, Mar. 2009.

[49] K. Grauman, M. Betke, J. Lombardi, J. Gips, and G. R. Bradski, "Communication via eye blinks and eyebrow raises: video-based human-computer interfaces," *Universal Access Inf. Soc.*, vol. 2, no. 4, pp. 359–373, Nov. 2003.

[50] V. Rantanen, P.-H. Niemenlehto, J. Verho, and J. Lekkala, "Capacitive facial movement detection for human–computer interaction to click by frowning and lifting eyebrows," *Med. Biol. Eng. Comput.*, vol. 48, no. 1, pp. 39–47, Jan. 2010.

[51] I. R. Titze and D. W. Martin, "Principles of voice production," *Acoust. Soc. Amer. J.*, vol. 104, p. 1148, Sep. 1998, doi: 10.1121/1.424266.

[52] T. Igarashi and J. F. Hughes, "Voice as sound: Using non-verbal voice input for interactive control," in *Proc. 14th Annu. ACM Symp. Interface Softw. Technol. (UIST)*, Orlando, FL, USA, 2001, pp. 155–156.

[53] S. Harada, J. A. Landay, J. Malkin, X. Li, and J. A. Bilmes, "The vocal joystick: Evaluation of voice-based cursor control techniques," in *Proc. 8th Int. ACM SIGACCESS Conf. Comput. Accessibility (Assets)*, Portland, 2006, pp. 197–204.

[54] J. Janer, "Singing-driven interfaces for sound synthesizers," Ph.D. dissertation, Dept. Inf. Commun. Technol., Pompeu Fabra Univ., Barcelona, Spain, 2008.

[55] D. Stowell, "Making music through real-time voice timbre analysis: Machine learning and Timbral control," Ph.D. dissertation, School Electron. Eng. Comput. Sci., Queen Mary Univ., London, U.K., 2010.

[56] S. Fasciani, "Voice-controlled interface for digital musical instruments," Ph.D. dissertation, Citeseer/Grad. School Integrative Sci. Eng., Nat. Univ. Singapore, Singapore, 2014.

[57] D. O'Shaughnessy, "Interacting with computers by voice: Automatic speech recognition and synthesis," *Proc. IEEE*, vol. 91, no. 9, pp. 1272–1305, Sep. 2003.

[58] T. A. Wilson, G. S. Beavers, M. A. DeCoster, D. K. Holger, and M. D. Regenfuss, "Experiments on the fluid mechanics of whistling," *J. Acoust. Soc. Amer.*, vol. 50, no. 1B, pp. 366–372, Jul. 1971.

[59] A. Azola, J. Palmer, R. Mulheren, R. Hofer, F. Fischmeister, and W. T. Fitch, "The physiology of oral whistling: A combined radiographic and MRI analysis," *J. Appl. Physiol.*, vol. 124, no. 1, pp. 34–39, Jan. 2018.

[60] U. Esnaola and T. Smithers, "Whistling to machines," in *Ambient Intelligence in Everyday Life*, Y. Cai and J. Abascal, Eds. Berlin, Germany: Springer-Verlag, 2006, pp. 198–226.

[61] W. Birmingham, R. Dannenberg, and B. Pardo, "Query by humming with the vocalsearch system," *Commun. ACM*, vol. 49, no. 8, pp. 49–52, 2006.

[62] H.-C. Shen and C.-N. Lee, "An interactive Whistle-to-Music composing system based on transcription, variation and chords generation," *Multimedia Tools Appl.*, vol. 53, no. 1, pp. 253–269, May 2011.

[63] I. Mougharbel, R. El-Hajj, H. Ghamlouch, and E. Monacelli, "Comparative study on different adaptation approaches concerning a sip and puff controller for a powered wheelchair," in *Proc. IEEE Sci. Inf. Conf.*, London, U.K., Oct. 2013, pp. 597–603.

[64] M. Jones, K. Grogg, J. Anschutz, and R. Fierman, "A Sip-and-Puff wireless remote control for the apple iPod," *Assistive Technol.*, vol. 20, no. 2, pp. 107–110, Jun. 2008.

[65] T. K. Capin, H. Noser, D. Thalmann, I. Sunday Pandzic, and N. M. Thalmann, "Virtual human representation and communication in VLNet," *IEEE Comput. Graph. Appl.*, vol. 17, no. 2, pp. 42–53, Mar./Apr. 1997.

[66] B. Denby, T. Schultz, K. Honda, T. Hueber, J. M. Gilbert, and J. S. Brumberg, "Silent speech interfaces," *Speech Commun.*, vol. 52, no. 4, pp. 270–287, Apr. 2010.

[67] D. Bonafede, L. A. Ludovico, and G. Presti, "A proposal for the interactive sonification of the human face," in *Proc. 2nd Int. Conf. Comput.-Hum. Interact. Res. Appl.*, 2018, pp. 163–169.

[68] C. Salem and S. Zhai, "An isometric tongue pointing device," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. (CHI)*, Atlanta, 1997, pp. 538–539.

[69] X. Huo, J. Wang, and M. Ghovanloo, "A magneto-inductive sensor based wireless tongue-computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 5, pp. 497–504, Oct. 2008.

[70] H. Park, M. Kiani, H.-M. Lee, J. Kim, J. Block, B. Gosselin, and M. Ghovanloo, "A wireless magnetoresistive sensing system for an intraoral tongue-computer interface," *IEEE Trans. Biomed. Circuits Syst.*, vol. 6, no. 6, pp. 571–585, Dec. 2012.

[71] L. N. S. A. Struijk, "An inductive tongue computer interface for control of computers and assistive devices," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2594–2597, Dec. 2006.

[72] R. Vaidyanathan, B. Chung, L. Gupta, H. Kook, S. Kota, and J. D. West, "Tongue-movement communication and control concept for hands-free Human–Machine interfaces," *IEEE Trans. Syst., Man, Cybern., A, Syst. Humans*, vol. 37, no. 4, pp. 533–546, Jul. 2007.

[73] J. Cheng, A. Okoso, K. Kunze, N. Henze, A. Schmidt, P. Lukowicz, and K. Kise, "On the tip of my tongue: A non-invasive pressure-based tongue interface," in *Proc. 5th Augmented Hum. Int. Conf. (AH)*, 2014, p. 12.

[74] D. Wessel, "Instruments that learn, refined controllers, and source model loudspeakers," *Comput. Music J.*, vol. 15, no. 4, pp. 82–86, 1991.

[75] F. Vogt, G. McCaig, M. A. Ali, and S. S. Fels, "Tongue'n'Groove: An ultrasound based music controller," in *Proc. Int. Conf. New Interface Musical Expression (NIME2002)*, Dublin, 2002, pp. 60–64.

[76] H. Y. Nam and C. DiSalvo, "Tongue music: The sound of a kiss," in *Proc. Extended Abstr. Hum. Factors Comput. Syst.*, Atlanta, Ga, USA, 2010, p. 4805.

[77] H. Jeong, J.-S. Kim, and J.-S. Choi, "A study of an EMG-controlled HCI method by clenching teeth," in *Proc. Asia–Pacific Conf. Comput. Hum. Interact.*, Rotorua, New Zealand, 2004, pp. 163–170.

[78] Y. Lee and M. Lee, "SMS application using EMG signal of clenching teeth for e-Health communication," *Telemed. e-Health*, vol. 14, no. 6, pp. 593–597, Aug. 2008.

[79] T. Simpson, C. Broughton, M. J. A. Gauthier, and A. Prochazka, "Tooth-click control of a hands-free computer interface," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 8, pp. 2050–2056, Aug. 2008.

[80] T. Vega Gálvez, S. Sapkota, A. Dancu, and P. Maes, "Byte.it: Discreet teeth gestures for mobile device interaction," in *Proc. ACM Conf. Hum. Factors Comput. Syst. (CHI)*, 2019, Art. no. LBW2310.

[81] K. Kuzume, "Evaluation of tooth-touch sound and expiration based mouse device for disabled persons," in *Proc. IEEE Int. Conf. Pervas. Comput. Commun. Workshops*, Mar. 2012, pp. 387–390.

[82] D. Ashbrook, C. Tejada, D. Mehta, A. Jiminez, G. Muralitharam, S. Gajendra, and R. Tallents, "Bitey: An exploration of tooth click gestures for hands-free user interface control," in *Proc. ACM Int. Conf. Hum.-Comput. Interact. Mobile Devices Services (MobileHCI)*, Florence, Italy, 2016, pp. 158–169.

[83] H. Iwata, H. Yano, T. Uemura, and T. Moriya, "Food simulator: A haptic interface for biting," in *Proc. IEEE Virtual Reality*, Chicago, IL, USA, Mar. 2004, pp. 51–57.

[84] L. M. Nashner and G. McCollum, "The organization of human postural movements: A formal basis and experimental synthesis," *Behav. Brain Sci.*, vol. 8, no. 1, pp. 135–150, Mar. 1985.

[85] B. Biguer, M. Jeannerod, and C. Prablanc, "The coordination of eye, head, and arm movements during reaching at a single visual target," *Exp. Brain Res.*, vol. 46, no. 2, pp. 301–304, May 1982.

[86] W. R. Thurlow and P. S. Runge, "Effect of induced head movements on localization of direction of sounds," *J. Acoust. Soc. Amer.*, vol. 42, no. 2, pp. 480–488, Aug. 1967.

[87] C. T. Ishi, H. Ishiguro, and N. Hagita, "Analysis of relationship between head motion events and speech in dialogue conversations," *Speech Commun.*, vol. 57, pp. 233–243, Feb. 2014.

[88] H. Sarig-Bahat, "Evidence for exercise therapy in mechanical neck disorders," *Manual Therapy*, vol. 8, no. 1, pp. 10–20, Feb. 2003.

[89] J. Leaman and H. M. La, "A comprehensive review of smart wheelchairs: Past, present, and future," *IEEE Trans. Human-Machine Syst.*, vol. 47, no. 4, pp. 486–499, Aug. 2017.

[90] S. Serafin, C. Erkut, J. Kojs, N. C. Nilsson, and R. Nordahl, "Virtual reality musical instruments: State of the art, design principles, and future directions," *Comput. Music J.*, vol. 40, no. 3, pp. 22–40, Sep. 2016.

[91] D. Merrill, "Head-tracking for gestural and continuous control of parameterized audio effects," in *Proc. Conf. New Interfaces Musical Expression(NIME)*, Montreal, QC, Canada, 2003, pp. 218–219.

[92] A. Kapur, A. J. Lazier, P. Davidson, R. S. Wilson, and P. R. Cook, "The electronic sitar controller," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, 2004, pp. 7–12.

[93] G. L. Hands and C. E. Stepp, "Effect of age on Human–Computer interface control via neck electromyography," *Interacting Comput.*, vol. 28, no. 1, pp. 47–54, Jan. 2016.

[94] T. O. Zander, C. Kothe, S. Jatzev, and M. Gaertner, "Enhancing human-computer interaction with input from active and passive brain-computer interfaces," in *Brain-Computer Interfaces–Applying Our Minds to Human-Computer Interaction*, D. S. Tan and A. Nijholt, Eds. London, U.K.: Springer-Verlag, 2010, pp. 181–199.

[95] E. R. Miranda and J. Castet, Eds., *Guide to Brain-Computer Music Interfacing*. London, U.K.: Springer-Verlag, 2014.

[96] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, Sep. 1970.

[97] M. Singh, M. Singh, and S. Gangwar, "Emotion recognition using electroencephalography (EEG): A review," *Int. J. Inf. Technol. Knowl. Manage.*, vol. 7, no. 1, pp. 1–5, Dec. 2013.

[98] C. Mühl, B. Allison, A. Nijholt, and G. Chanel, "A survey of affective brain computer interfaces: Principles, state-of-the-art, and challenges," *Brain-Comput. Interfaces*, vol. 1, no. 2, pp. 66–84, Apr. 2014.

[99] J. A. Russell, "A circumplex model of affect," *J. Personality Social Psychol.*, vol. 39, no. 6, pp. 1161–1178, Dec. 1980.

[100] B. F. Yuksel, K. B. Oleson, L. Harrison, E. M. Peck, D. Afergan, R. Chang, and R. J. Jacob, "Learn piano with BACh: An adaptive learning interface that adjusts task difficulty based on brain state," in *Proc. ACM Conf. Hum. Factors Comput. Syst. (CHI)*, San Jose, CA, USA, 2016, pp. 5372–5384.

[101] B. F. Yuksel, D. Afergan, E. Peck, G. Griffin, L. Harrison, N. Chen, R. Chang, and R. Jacob, "BRAAHMS: A novel adaptive musical interface based on users' cognitive state," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, Baton Rouge, LA, USA, 2015, pp. 136–139.

[102] F. Morreale and A. De Angeli, "Collaborating with an autonomous agent to generate affective music," *Comput. Entertainment*, vol. 14, no. 3, pp. 1–21, Jan. 2017.

[103] A. Girouard, E. T. Solovey, and R. J. Jacob, "Designing a passive brain computer interface using real time classification of functional near-infrared spectroscopy," *Int. J. Auton. Adapt. Commun. Syst.*, vol. 6, no. 1, pp. 26–44, 2013.

[104] I. Daly, D. Williams, A. Kirke, J. Weaver, A. Malik, F. Hwang, E. Miranda, and S. J. Nasuto, "Affective brain–computer music interfacing," *J. Neural Engin.*, vol. 13, no. 4, 2016, Art. no. 046022.

[105] A. Väljamäe, T. Steffert, S. Holland, X. Marimon, R. Benitez, S. Mealla, A. Oliveira, and S. Jordà, "A review of real-time EEG sonification research," in *Proc. Int. Conf. Auditory Display*, 2013, pp. 85–93.

[106] F. Aloise, P. Aricò, F. Schettini, A. Riccio, M. Risetti, S. Salinari, D. Mattia, F. Babiloni, and F. Cincotti, "A new P300 no eye-gaze based interface: GeoSpell," in *Proc. Int. Conf. Bio-Inspired Syst. Signal Process. (BIOSIGNALS)*. Rome, Italy: SciTePress-Science and and Technology Publications, 2011, pp. 227–232.

[107] E. Donchin, K. M. Spencer, and R. Wijesinghe, "The mental prosthesis: Assessing the speed of a P300-based brain-computer interface," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 174–179, Jun. 2000.

[108] D. Zhu, J. Bieger, G. Garcia Molina, and R. M. Aarts, "A survey of stimulation methods used in SSVEP-based BCIs," *Comput. Intell. Neurosci.*, vol. 2010, pp. 1–12, 2010.

[109] R. Chavarriaga and J. del R. Millán, "Learning from EEG error-related potentials in noninvasive brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 18, no. 4, pp. 381–388, Aug. 2010.

[110] M. Grierson, "Composing with brainwaves: Minimal trial P300 recognition as an indication of subjective preference for the control of a musical instrument," in *Proc. Int. Comput. Music Conf. (ICMC)*, Belfast, Northern Ireland, 2008, pp. 1–4.

[111] A. Pinegger, H. Hiebel, S. C. Wriessnegger, and G. R. Müller-Putz, "Composing only by thought: Novel application of the P300 brain-computer interface," *PLoS ONE*, vol. 12, no. 9, Sep. 2017, Art. no. e0181584.

[112] Y. C. D. Chew and E. Caspary, "MusEEGk: A brain computer musical interface," in *Proc. Extended Abstr. Hum. Factors Comput. Syst.*, Vancouver, BC, Canada, 2011, p. 1417.

[113] E. R. Miranda,"Brain-computer music interface for composition and performance," *Int. J. Disability Hum. Develop.*, vol. 5, no. 2, p. 119, Jan. 2006.

[114] E. Miranda and A. Brouse, "Toward direct brain-computer musical interfaces," in *Proc. Conf. New Interfaces Musical Expression (NIME)*, Vancouver, BC, Canada, May 2005, p. 4.

[115] E. V. Friedrich, R. Scherer, and C. Neuper, "The effect of distinct mental strategies on classification performance for brain–computer interfaces," *Int. J. Psychophysiol.*, vol. 84, no. 1, pp. 86–94, 2012.

[116] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt, and J. P. Donoghue, "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, May 2012.

[117] K. LaFleur, K. Cassady, A. Doud, K. Shades, E. Rogin, and B. He, "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface," *J. Neural Eng.*, vol. 10, no. 4, Aug. 2013, Art. no. 046003.

[118] N. Birbaumer, A. Kubler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, and H. Flor, "The thought translation device (TTD) for completely paralyzed patients," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 190–193, Jun. 2000.

[119] M. Pham, T. Hinterberger, N. Neumann, A. Kübler, N. Hofmayer, A. Grether, B. Wilhelm, J.-J. Vatine, and N. Birbaumer, "An auditory brain-computer interface based on the self-regulation of slow cortical potentials," *Neurorehabil. Neural Repair*, vol. 19, no. 3, pp. 206–218, Sep. 2005.

[120] Z. Vamvakousis and R. Ramirez, "Towards a low cost mu-rhythm based BCI," in *Proc. Int. Brain-Comput. Interface Meeting*, Pacific Grove, CA, USA, 2013, p. 13.

[121] T. A. Deuel, J. Pampin, J. Sundstrom, and F. Darvas, "The encephalophone: A novel musical biofeedback device using conscious control of electroencephalogram (EEG)," *Frontiers Hum. Neurosci.*, vol. 11, P. 213, Apr. 2017.

[122] T. Hamano, T. M. Rutkowski, H. Terasawa, K. Okanoya, and K. Furukawa, "Generating an integrated musical expression with a brain-computer interface," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, 2013, pp. 49–54.

[123] R. B. Knapp and H. S. Lusted, "A bioelectric controller for computer music applications," *Comput. Music J.*, vol. 14, no. 1, pp. 42–47, 1990.

[124] E. Lyon, R. B. Knapp, and G. Ouzounian, "Compositional and performance mapping in computer chamber music: A case study," *Comput. Music J.*, vol. 38, no. 3, pp. 64–75, Sep. 2014.

[125] T. Niikawa, "Tongue-controlled electro-musical instrument," in *Proc. 18th Int. Congr. Acoust. (ICA)*, Kyoto, Japan, Apr. 2004, pp. 1905–1909.

[126] V. Matossian and R. Gehlhaar, "Human instruments: Accessible musical instruments for people with varied physical ability," *Annu. Rev. Cybertherapy Telemed.*, vol. 13, pp. 200–205, 2015.

[127] HumanInstruments. *Instruments*. Accessed: Dec. 7, 2019. [Online]. Available: https://www.humaninstruments.co.uk/n.d

[128] Housemate. *Magic Flute*. Accessed: Jun. 8, 2019. [Online]. Available: http://housemate.ie/magic-flute/n.d

[129] Jamboxx. *Jamboxx*. Accessed: Jun. 8, 2019. [Online]. Available: https://www.jamboxx.com/

[130] S. Bailey, A. Scott, H. Wright, I. Symonds, and K. Ng, "Eye. Breathe. music: A multimeadia controller that uses eye movement and breathing," in *Proc. Int. Comput. Music Conf. (ICMC)*, Ann Arbor, MI, USA, 2010, pp. 104–107.

[131] E. R. Miranda, W. L. Magee, J. J. Wilson, J. Eaton, and R. Palaniappan, "Brain-computer music interfacing (BCMI): From basic research to the real world of special needs," *Music Med.*, vol. 3, no. 3, pp. 134–140, Jul. 2011.

[132] J. Eaton, "Brain-computer music interfacing: Designing practical systems for creative applications," Ph.D. dissertation, School Humanities Performing Arts, Interdiscipl. Centre Comput. Music Res., Fac. Arts, Univ. Plymouth, Plymouth, U.K., 2016.

[133] Z. Vamvakousis and R. Ramirez, "Temporal control in the eyeharp gaze-controlled musical interface," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, Ann Arbor, MI, USA, 2012, pp. 1–16.

[134] Z. Vamvakousis and R. Ramirez, "The EyeHarp: A gaze-controlled digital musical instrument," *Frontiers Psychol.*, vol. 7, P. 9067, Jun. 2016.

[135] ——, "A Brain-Gaze Controlled Musical Interface," in *Proc. Berlin BCI Workshop Adv. Neurotechnol.*, Berlin, Germany, Sep. 2012, p. 2.

[136] Fove. *Eye Play the Piano*. Accessed: Jun. 8, 2019. [Online]. Available: http://eyeplaythepiano.com/en/

[137] Z. Vamvakousis and R. Ramirez, "P300 Harmonies: A Brain-Computer Musical Interface," in *Proc. Int. Comput. Music Conf. (ICMC)*, Athens, Greece, Sep. 2014, p. 5.

[138] Imitone. *Imitone: Mind to Melody*. Accessed: Jun. 8, 2019. [Online]. Available: https://imitone.com/

[139] OpenUpMusic. *The Clarion*. Accessed: Dec. 7, 2019. [Online]. Available: http://openupmusic.org/the-clarion/

[140] C. H. Morimoto, A. Diaz-Tula, J. A. T. Leyva, and C. E. L. Elmadjian, "Eyejam: A gaze-controlled musical interface," in *Proc. 14th Brazilian Symp. Hum. Factors Comput. Syst. (IHC)*, New York, NY, USA, 2015, p. 37.

[141] A. Refsgaard. *Eye Conductor*. Accessed: Jun. 8, 2019. [Online]. Available: https://andreasrefsgaard.dk/ project/eye-conductor/

[142] N. Davanzo, P. Dondi, M. Mosconi, and M. Porta, "Playing music with the eyes through an isomorphic interface," in *Proc. Worksh. Commun. Gaze Interact. (COGAIN)*, Warsaw, Poland, 2018, pp. 1–5.

[143] J. Vogels, "Harmonica-inspired digital musical instrument design based on an existing gestural performance repertoire," M.S. thesis, Music Technol. Area, Schulich School Music, McGill Univ., Montreal, QC, Canada, 2015.

[144] R. Fiebrink, D. Trueman, and P. R. Cook, "A meta-instrument for interactive, on-the-fly machine learning," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, Pittsburgh, PA, USA, 2009, pp. 280–285.

[145] S. Katan, M. Grierson, and R. Fiebrink, "Using interactive machine learning to support interface development through workshops with disabled people," in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst. (CHI)*, 2015, pp. 251–254.

[146] L. Turchet, "Smart musical instruments: Vision, design principles, and future directions," *IEEE Access*, vol. 7, pp. 8944–8963, 2019.

[147] A. Marquez-Borbon and J. P. Martinez Avila, "The problem of DMI adoption and longevity: Envisioning a NIME performance pedagogy," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, Blacksburg, VA, USA, 2018, pp. 190–195.

[148] D. Schuler and A. Namioka, Eds., *Participatory Design: Principles and Practices*. Boca Raton, FL, USA: CRC Press, 1993.

[149] A. Marquez-Borbon and P. Stapleton, "Fourteen years of NIME: the value and meaning of 'community' in interactive music research," in *Proc. Int. Conf. New Interfaces Musical Expression (NIME)*, Baton Rouge, LA, USA, 2015, pp. 307–312.

NICOLA DAVANZO received the M.Sc. degree in computer engineering from the University of Pavia, in 2018, with a thesis titled Development and testing of graphical interfaces for eye-controlled musical instruments. He is currently pursuing the Ph.D. degree in computer science with the University of Milan. His current research project aims the design and development of accessible digital musical instruments, with a particular focus on upper limbs physical disabilities. His research interests include HCI and music in a general sense.

FEDERICO AVANZINI received the Ph.D. degree in computer science from the University of Padova, in 2002. Until 2017, he was a Postdoctoral Researcher, an Assistant Professor, and an Associate Professor with the University of Padova. He is currently an Associate Professor with the University of Milan.

He has authored about 180 publications on peer-reviewed international journals and conferences. His current research interests include concern algorithms for sound synthesis and processing, non-speech sound in human–computer interfaces, and multimodal interaction. He has been a key researcher and a principal investigator in several national and international research projects. He has served for several program and editorial committees. He has chaired international conferences and workshops. He is currently an Associate Editor of the international journal *Acta Acustica*. He is also the President of the Italian Music Informatics Association.

• • •