

Introduction to Intelligent Music Systems and Applications

MARKUS SCHEDL, Johannes Kepler University, Linz, Austria

YI-HSUAN YANG, Academia Sinica, Taipei, Taiwan

PERFECTO HERRERA-BOYER, Universitat Pompeu Fabra and Escola Superior de Música de Catalunya, Barcelona, Spain

Intelligent technologies have become an essential part of music systems and applications. This is evidenced by today's omnipresence of digital online music stores and streaming services, which rely on music recommenders, automatic playlist generators, and music browsing interfaces. A large amount of research leading to intelligent music applications deals with the extraction of musical and acoustic information directly from the audio signal using signal processing techniques. Other strategies exploit contextual aspects of music, not present in the signal, for example, community meta-data and trails of user interaction, as found, for instance, on social media platforms. In this editorial, we discuss the notion of "intelligent music system" and give an overview of the papers selected to this special issue.

CCS Concepts: • **Applied computing** → **Sound and music computing**

Additional Key Words and Phrases: Intelligent music systems, machine learning, music information retrieval

ACM Reference Format:

Markus Schedl, Yi-Hsuan Yang, and Perfecto Herrera-Boyer. 2016. Introduction to intelligent music systems and applications. *ACM Trans. Intell. Syst. Technol.* 8, 2, Article 17 (October 2016), 8 pages.

DOI: <http://dx.doi.org/10.1145/2991468>

1. THE CHALLENGE TO DEFINE WHAT "INTELLIGENT MUSIC SYSTEMS" ARE

There are many devices, applications, and systems around us that are termed "intelligent." Being more specifically related to music, we can compose music with intelligent arrangers,^{1,2} play intelligent instruments³ while we process their sounds through intelligent sound effects,⁴ play along with intelligent improvisers,⁵ browse/play music collections by means of intelligent interfaces,⁶ or engage with games that adapt the soundtrack in real time to the current characters and action.⁷

Being intelligent is, and probably will always be, "cool." Therefore, researchers and developers strive to have reasons to add the "intelligent" adjective to their systems. Intelligence is a controversial concept even though we adopt a pragmatic, computational, empirical, or quantitative view and try to avoid the more obscure epistemological

¹<http://www.ntonyx.com>.

²<http://www.pgmusic.com/bbwin.htm>.

³<http://www.youtube.com/watch?v=D-mmEvGOopk>.

⁴<https://www.izotope.com/en/products/mix/nectar/features.html>.

⁵<http://www.wired.com/2008/11/no-way-robot-ja>.

⁶<http://www.bang-olufsen.com/en/sound/sound-systems/beosound-moment>.

⁷<http://www.intelligentmusicsystems.com>.

Authors' addresses: M. Schedl, Johannes Kepler University Linz, Altenberger Straße 69, 4040 Linz, Austria; email: markus.schedl@jku.at; Y.-H. Yang, 128 Academia Road, Section 2, Nankang, Taipei 11564, Taiwan; email: yang@citi.sinica.edu.tw; P. Herrera-Boyer, Universitat Pompeu Fabra, Roc Boronat, 138, 08018 Barcelona, Spain; email: perfecto.herrera@upf.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

2016 Copyright is held by the owner/author(s).

2157-6904/2016/10-ART17

DOI: <http://dx.doi.org/10.1145/2991468>

or philosophical perspectives. Some researchers [Legg and Hutter 2007] have listed and discussed many definitions (more than 60!) proposed by psychologists, computer scientists, educators, mathematicians, and so forth, suggesting finally this synthetic definition: “intelligence measures an agent’s ability to achieve goals in a wide range of environments.” This is a definition of general intelligence that could probably be streamlined to fit our music domain, but it would still leave more issues open than clarifying them. Although some authors claim that “until we have a more precise definition of intelligence, the quest for generally intelligent machines will lack reliable techniques for measuring progress” [Legg and Veness 2013], there are many others that avoid such a definition. In fact, this has been the case for many of the authors published in this special issue, who were explicitly asked to provide, in their contributions, some definition and discussion on the concept. Provokingly said, it seems that doing research on intelligent systems can be a hot topic, provided we keep the concept undefined or vague. We contend here that intelligence, as beauty, is mostly in the eyes of the beholder. We can measure properties of the system’s behavior and, provided we have beforehand set some “intelligence traces,” we can quantitatively estimate intelligence to some extent. Let us remark that we need the investigated system to show some behavior. This means that intelligence cannot be determined by examining the code or inner working of the system. It is therefore a functional property of systems—something that is not implementation dependent nor a specific computational goal, to follow David Marr’s proposed levels [Marr 1982]. It is by means of action, by observing the behavior of the system that we can decide on its intelligence. Additionally, intelligence is not a Boolean property but a matter of degree. Degrees of intelligence, as they can be found when different animal species are studied and tested in this respect (i.e., a lizard is considered to be less intelligent than a pigeon, a pigeon less than a parrot, etc.), can also be considered for quantifying computing systems. Some researchers [Gardner 1993] have proposed that such a general aptitude is, in fact, a combination of several specializations (i.e., “multiple intelligences” dealing with verbal, logical, spatial, naturalistic, kinesthetic, interpersonal, and musical competence). Musical intelligence could either be a capability to build interconnections among the different layers of a musical work (rhythm, melody, harmony, timbre, etc.), to generate sequences that other systems consider as “music,” or to anticipate musical features or events while listening to a musical performance, or all of them at the same time. Musical intelligence, then, qualifies for specific approaches and methodologies, in accordance with the addressed problem. This also applies to the so-called intelligent music systems (IMSs), which can range from automatic composition systems that are capable of showing some form of creativity, music recommendation systems that provide justified and personalized reasons to suggest a track, “virtual DJs” that create dance-music sessions by combining and manipulating tracks, or automatic mixers that take a multitrack recording and turn them into a good-sounding, well-produced stereo track.

2. EXAMPLES OF INTELLIGENT MUSIC SYSTEMS

Given the sluggishness of the concept and our nonexhaustive purpose, we do not aim to provide a state of the art of IMSs, but just to mention some examples that we deem interesting. In addition to these, the interested reader may find valuable information in some early collections of papers and articles.^{8,9}

One category of IMSs is systems for automatic music composition and creation, such as the ones that have been proposed and described by Cope since 1981 in several books (e.g., Cope [1996, 2004, 2005]). His algorithms have since then produced

⁸<http://www.springeropen.com/collections/iasmpa>.

⁹<http://www.igi-global.com/book/intelligent-music-information-systems/633>.

classical music ranging from single-instrument arrangements all the way up to full symphonies by modeling the styles of composers like Bach, Mozart, or Stravinsky, or even “inventing” specific styles from virtual composers (e.g., “Emily Howell”). On many occasions, listeners have been fooled into believing that the works were written by human composers, which has made some people consider some of them as clear examples of passing a Turing test. Experiments in musical intelligence are based on three basic principles: deconstruction of parts, commonality (or style signature) detection, and recombination of parts into new works that respect style signatures. More recent “virtual composers” that have called for attention are *Iamus*¹⁰ and *Melomisc109*,¹¹ based on genetic algorithms that encode melodies indirectly and on “epigenetic” processes that process those codes to generate songs [Delgado et al. 2009]. Another group of generative IMSs requires some human input as a basis to harmonize or improvise. For instance, *CHORAL* [Ebcioğlu 1990] is an expert system for harmonizing four-part chorales in the style of Bach. The system uses several hundred rules, expressed in a form of first-order predicate calculus, for representing the knowledge required for harmonizing a given melody. The rules observe the chorale from multiple viewpoints, such as the chord skeleton, individual melodic lines of each voice, and voice leading. The program harmonizes chorales using a generate-and-test method with intelligent backtracking. Charles Ames’s *Cybernetic Composer* [Ames and Domino 1992] creates music in popular and jazz styles based on basic chord progressions. Whether composing rock or ragtime, *Cybernetic Composer* often produces quite musical results. Pachet’s *Continuator* combines music interaction and composition [Pachet 2002, 2004]. It learns in real time stylistic features of a music player and interactively plays with him or her in that style. It is based on an augmented Markov model of musical styles. The system builds operational representations of musical styles in real time and quickly adapts to external musical information. In a similar way, but based on different formal principles and software tools, *OMax* [Assayag and Dubnov 2004; Assayag et al. 2006] learns a musician’s style and co-improvises an accompaniment, adapting its music to what is happening in the performance. *OMax* uses the so-called stylistic reinjection and a semantic representation level of the session in order to recombine and transform materials in real time.

Two examples of IMSs, which are quite different from the previously described systems, offer personalized and context-aware music recommendations: *Just-for-Me* [Cheng and Shen 2014] and *Mobile Music Genius* [Schedl et al. 2014]. The former builds a model based on the listener’s short-term and long-term music preferences, his or her location, the general popularity of music items, and audio features. This information is used to train a probabilistic latent topic model that in turn allows one to estimate in which context the user will likely want to listen to which track and to suggest respective music accordingly. The latter system provides a means for automatic playlist adaptation on mobile devices. To this end, the user’s interaction with the music player, as well as features ranging from weather to environmental noise and motion data, are continuously monitored. A classifier is trained on these features and in turn used to predict in which context the listener presumably likes a certain music item. If the user context substantially changes, the playlist is modified according to this prediction.

Another category of systems that we consider IMSs are applications that provide interactive ways to browse or explore music collections. One example of such a system is the *Musicream*¹² interface that aims to foster serendipitous music discovery [Goto

¹⁰<http://melomics.com/iamus>.

¹¹<http://melomics.com/melomics109>.

¹²<https://staff.aist.go.jp/m.goto/Musicream>.

and Goto 2010]. Such systems automatically organize music repositories according to some meaningful criteria, typically inferred from low-level audio representations (e.g., acoustic similarity, mood, or style), present the resulting organization to the users, and let them interact with the collection in intuitive ways. To this end, Musicream employs the metaphor of water taps that release streams of music pieces, represented as discs, when the user opens the respective faucets. Different taps release music of different styles. The user can then grab pieces and listen to them, or create playlists by sticking together tracks or entire playlists. When doing so, pieces sounding alike are easier to connect than dissimilar ones, based on a physical model of attractive and repellent forces, respectively. Other interfaces employ clustering techniques to organize music collections according to some notion of similarity, which is inferred from the audio or from web information. The clusters are then visualized and interaction capabilities are provided to the users of such interfaces, examples of which are *nepTune*¹³ [Knees et al. 2006], *MusicGalaxy* [Stober and Nürnberger 2011], and an interface presented in Lübbers and Jarke [2009].

Recent years have also witnessed a growing interest in building algorithms and systems for predicting the emotion a listener may perceive in a piece of music, enabling automatic emotion-based organization and access to music information [Yang and Chen 2011]. State-of-the-art algorithms allow a user to create a music playlist of desired emotions by specifying a point or drawing a trajectory in a low-dimensional emotion space. Personalization techniques that learn from a user's feedback can be further employed to account for the subjective nature of emotion perception [Wang et al. 2016]. It is also possible to actively and continuously detect a listener's emotional state from physiological, facial, or textual cues and to recommend music in a personalized and "mood-optimizing" way [Yang and Chen 2012; Ferwerda and Schedl 2014].

In conclusion, it seems that most of what we tend to consider as "intelligent" exhibits observable behavior, such as musical/listenable output, and it is capable of interacting with (and adapting to) a changing environment, for instance, a composer, a partner that plays music, or a listener doing some activity. Although purely analytic approaches without any adaptive or behavioral component, for instance, an algorithm to extract the melody line from a polyphonic audio file, could not qualify as full-fledged intelligent, they are nevertheless crucial to build IMSs on top of them, which is reflected in the selection of articles for this special issue.

3. SELECTION PROCESS OF ARTICLES FOR THIS SPECIAL ISSUE

The call for papers of this special issue was perceived very positively, and we eventually received 38 regular submissions from which we selected five for inclusion, in a three-stage reviewing process. The acceptance rate was therefore 13%. In addition to regular papers, we further invited highly recognized experts in the field to submit thought-provoking overview or position articles. We selected two, namely, one by Gerhard Widmer and one by François Pachet, both receivers of the prestigious ERC Advanced Grant. All articles are briefly presented in the following.

4. CONTENT OF THIS SPECIAL ISSUE

The invited article by Widmer, *Getting Closer to the Essence of Music: The Con Espresione Manifesto*, provides an excellent overview about the current state of research in building systems with a deeper understanding of music than current ones we have. In addition, it highlights major problems still requiring intensive research, such as perception and appreciation of music, music performance description, long-term modeling of music sequences, or expressive and affective content characterization.

¹³<http://www.cp.jku.at/projects/neptune>.

Pachet, in our second guest article, *A Joyful Ode to Automatic Orchestration*, illustrates the process of semiautomatically reorchestrating Beethoven's Ode to Joy according to stylistic conventions of seven different musical models (e.g., Bach chorales, Ennio Morricone soundtracks, jazz, or bossa nova). The article helps to gauge where the limits of several current state-of-the-art techniques are, and how and where musical knowledge has to be manually added to come up with a final music production (this becoming then an implicit roadmap for open research issues). The article also highlights our shortcomings in modeling the interplay between expressiveness, timbre, and symbolic content, especially when creating pop music.

Among the regular articles, the contribution by Sandouk and Chen, *Learning Contextualized Music Semantics from Tags via a Siamese Neural Network*, investigates the use of a Siamese neural network architecture, based on bag-of-words representations of tags and on deep learning, for two tasks related to automatic music annotation, namely, semantic priming and tag completion. Such techniques to automatically infer semantic knowledge are vital building blocks for many IMSs, including intelligent music interaction interfaces and music emotion recognition and regulation. The authors show (on three popular music tag datasets, CAL500, MagTak5K, and a subset of the Million Song Dataset) that the semantics learned by their concept embedding model outperforms that of other state-of-the-art approaches, both qualitatively by visualization and quantitatively by common IR measures. This is particularly the case when tags are scarcely available.

Schindler and Rauber, in *Harnessing Music-Related Visual Stereotypes for Music Information Retrieval*, advocate exploiting music-related visual material, such as music videos, album covers, and advertisements in music retrieval and recommendation systems. The rationale is that such an approach may capture information that is not available in the audio signal. As a proof of concept, the authors consider a task that uses low- to mid-level audio and visual features extracted from music videos to assign high-level semantic labels, such as genres and themes, to the music. Good solutions to this task of learning semantic concepts from various representations of music, while not considered IMSs themselves, are essential building blocks for them. State-of-the-art visual concept detectors based on deep learning are also employed for feature extraction. The experiments confirm the effectiveness of combining the audio and visual modalities. The labels and extracted features of the newly created Music Video Dataset are made publicly available to facilitate research along this line.

Oramas et al., in *Sound and Music Recommendation with Knowledge Graphs*, explore a knowledge-graph-based approach to supply a recommendation engine with information computed from web documents describing musical and sound items. Specifically, entity-linking techniques are employed to extract semantic entities from item textual descriptions and to link them to external knowledge graphs, such as WordNet and DBpedia, to exploit additional knowledge. Two graph feature mappings are used to leverage the resulting knowledge graph for making recommendations. Quantitative experiments, using sounds from FreeSound.org and tracks from Last.fm and Songfacts.com, show that this approach leads to improved accuracy, novelty, and diversity compared to existing collaborative filtering and content-based recommendation algorithms. Sophisticated music recommender systems that interact with their users could be considered as a primitive form of IMSs, notwithstanding that their behavior is restricted to updating a recommendation list based on user feedback and interaction traces.

Tian and Sandler, in *Towards Music Structural Segmentation Across Genres: Features, Structural Hypotheses, and Annotation Principles*, address the problem of automatic segmentation dealing with a type of Chinese music called Jingju. Several Western-centric biases of many of the existing algorithms and concepts used for the

automatic analysis and description of music are illustrated (and overcome) here. Instead of basing decisions on tonality and recurrence notions, thanks to the adoption of a musicological perspective, the proposed system capitalizes its improved performance on timbre and novelty features. The problem of automatic or autonomous feature selection, even though not explicitly touched on here, arises as one of the required components for many IMSs.

Carabias-Orti et al., in *Tempo Driven Audio-to-Score Alignment Using Spectral Decomposition and Online Dynamic Time Warping*, propose a new algorithm for aligning a MIDI score with its audio performance in an online scenario where the data has to be processed as the audio signal is acquired. This is a core technique in interactive IMSs that involves a computer “listening” and responding to a performer’s audio input in real time, for example, by turning the pages of the music score or by accompanying the performance with automatically generated music. The new algorithm uses non-negative matrix factorization (NMF) to measure the similarity between a synthesized MIDI signal and the input audio signal, and employs robust variants of online dynamic time warping (DTW) for alignment. The algorithm is shown to be more accurate than other state-of-the-art methods in the literature. The article gives a good example of connecting the symbolic and audio information of music in IMSs.

5. CONCLUDING REMARKS

It seems difficult to agree on what intelligence of a computer system is, how it can be examined or measured, or even whether the question of evaluating the intelligence of a system should be considered a research problem. We should be cautious on the abuse of the term “intelligence” in order to preserve its meaning and power but, at the same time, be open to acknowledge the possibilities of a clever computation or a powerful algorithm that might give way to intelligent behavior when combined with others in a larger system. The articles included here would not pass the well-known Turing test [Turing 1950], but they are relevant contributions that might pave the way to intelligent full-fledged systems. The pervasiveness of the Turing test (see, for example, the recent *AlgoRhythms* contest¹⁴) should not conceal its limitations and hence there is a need to consider other perspectives that the music computing community has not tried yet. Winograd schemes [Levesque 2011] have been recently used, but they rely on language processing. Other formal proposals do exist [Legg and Veness 2013], for example, the Lovelace test [Bringsjord et al. 2001] or alternative settings in which Turing-like tests should be carried out.¹⁵ Even workshops on the topic have been held, for example, the Beyond the Turing Test workshop¹⁶ in conjunction with the AAAI Conference on Artificial Intelligence in 2015.

The composition of this special issue certainly represents only a snapshot of current topics and approaches toward building intelligent music systems and applications. It is by no means exhaustive, nor does it cover all areas of what we believe such systems and applications are. Much to our regret, we had to exclude many submissions of high quality due to space limitations. We nevertheless believe to have assembled a highly interesting assortment of state-of-the-art research and position articles, and we sincerely hope that you will enjoy reading this special issue as much as we enjoyed curating it.

The guest editors of this special issue.

¹⁴<http://bregman.dartmouth.edu/turingtests>.

¹⁵<http://spectrum.ieee.org/robotics/artificial-intelligence/artificialintelligence-experts-to-explore-turing-test-triathlon>.

¹⁶<http://www.aaai.org/Workshops/ws15workshops.php#ws06>.

ACKNOWLEDGMENTS

We would like to thank all authors for their highly interesting contributions, the reviewers for their incredible dedication to the special issue, and the editor-in-chief and his assistant for their support. The special issue further received support by the Austrian Science Fund (FWF): P25655, the Ministry of Science and Technology of Taiwan (MOST102-2221-E-001-004-MY3), and the EU-funded GiantSteps project (FP7-ICT-2013-10, grant agreement no. 610591).

REFERENCES

- Charles Ames and Michael Domino. 1992. Cybernetic composer: An overview. In *Understanding Music with AI*. MIT Press, Cambridge, MA, 186–205.
- G rard Assayag, Georges Bloch, Marc Chemillier, Arshia Cont, and Shlomo Dubnov. 2006. OMax brothers: A dynamic topology of agents for improvisation learning. In *Proceedings of the ACM Workshop on Audio and Music Computing Multimedia*. 125–132.
- G rard Assayag and Shlomo Dubnov. 2004. Using factor oracles for machine improvisation. *Soft Computing* 8, 9 (2004), 604–610. DOI: <http://dx.doi.org/10.1007/s00500-004-0385-4>
- Selmer Bringsjord, Paul Bello, and David Ferrucci. 2001. Creativity, the Turing test, and the (better) Lovelace test. *Minds and Machines* 11, 1 (2001), 3–27. DOI: <http://dx.doi.org/10.1023/A:1011206622741>
- Zhiyong Cheng and Jialie Shen. 2014. Just-for-me: An adaptive personalization system for location-aware social music recommendation. In *Proceedings of the ACM International Conference on Multimedia Retrieval*. 1267–1268.
- David Cope. 1996. *Experiments in Musical Intelligence*. A-R Editions, Madison, WI.
- David Cope. 2004. *Virtual Music: Computer Synthesis of Musical Style*. MIT Press, Cambridge, MA.
- David Cope. 2005. *Computer Models of Musical Creativity*. MIT Press, Cambridge, MA.
- Miguel Delgado, Waldo Fajardo, and Miguel Molina-Solana. 2009. Inmamusys: Intelligent multiagent music system. *Expert Systems with Applications* 36, 3 (2009), 4574–4580. DOI: <http://dx.doi.org/10.1016/j.eswa.2008.05.028>
- Kemal Ebcioglu. 1990. An expert system for harmonizing chorales in the style of J. S. Bach. *Journal of Logic Programming — Special Issue: Logic Programming Applications* 8, 1–2 (1990), 145–185. DOI: [http://dx.doi.org/10.1016/0743-1066\(90\)90055-A](http://dx.doi.org/10.1016/0743-1066(90)90055-A)
- Bruce Ferwerda and Markus Schedl. 2014. Enhancing music recommender systems with personality information and emotional states: A proposal. In *Proceedings of the Conference on User Modeling, Adaptation, and Personalization*.
- Howard Gardner. 1993. *Frames of Mind: The Theory of Multiple Intelligences*. Basic Books, New York, NY.
- Masataka Goto and Takayuki Goto. 2010. Musicream: Integrated music-listening interface for active, flexible, and unexpected encounters with musical pieces. *Information and Media Technologies* 5, 1 (2010), 139–152. DOI: <http://dx.doi.org/10.11185/imt.5.139>
- Peter Knees, Markus Schedl, Tim Pohle, and Gerhard Widmer. 2006. An innovative three-dimensional user interface for exploring music collections enriched with meta-information from the web. In *Proceedings of ACM Multimedia*. 17–24. DOI: <http://dx.doi.org/10.1145/1180639.1180652>
- Shane Legg and Marcus Hutter. 2007. A collection of definitions of intelligence. In *Proceedings of the Conference on Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms*. 17–24.
- Shane Legg and Joel Veness. 2013. *An Approximation of the Universal Intelligence Measure*. Springer, Berlin, 236–249. DOI: http://dx.doi.org/10.1007/978-3-642-44958-1_18
- Hector J. Levesque. 2011. The Winograd schema challenge. In *Proceedings of the AAI Spring Symposium: Logical Formalizations of Commonsense Reasoning*.
- Dominik L ubbers and Matthias Jarke. 2009. Adaptive multimodal exploration of music collections. In *Proceedings of the 10th International Society for Music Information Retrieval Conference*. 195–200.
- David Marr. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Henry Holt and Co., New York, NY.
- Fran ois Pachet. 2002. Interacting with a musical learning system: The continuator. In *Proceedings of the International Conference on Music and Artificial Intelligence*. Springer, London, UK, 119–132.
- Fran ois Pachet. 2004. Beyond the cybernetic jam fantasy: The continuator. *IEEE Computer Graphics and Applications* 24, 1 (2004), 31–35. DOI: <http://dx.doi.org/10.1109/MCG.2004.1255806>
- Markus Schedl, Georg Breitschopf, and Bogdan Ionescu. 2014. Mobile music genius: Reggae at the beach, metal on a friday night? In *Proceedings of the ACM International Conference on Multimedia Retrieval*. 507–510.

- Sebastian Stober and Andreas Nürnberger. 2011. MusicGalaxy: A multi-focus zoomable interface for multi-facet exploration of music collections. In *Exploring Music Contents: 7th International Symposium (CMMR'10)*. Málaga, Spain, June 21–24, 2010. Revised Papers. Springer, Heidelberg, Germany, 293–302. DOI: http://dx.doi.org/10.1007/978-3-642-23126-1_18
- Alan M. Turing. 1950. Computing machinery and intelligence. *Mind* 59, 236 (1950), 433–460. DOI: <http://dx.doi.org/10.1093/mind/LIX.236.433>
- Ju-Chiang Wang, Yi-Hsuan Yang, and Hsin-Min Wang. 2016. Affective music information retrieval. In *Emotions and Personality in Personalized Services: Models, Evaluation and Applications*, Marko Tkaličič et al. (Eds.). Springer.
- Yi-Hsuan Yang and Homer H. Chen. 2011. *Music Emotion Recognition*. CRC Press, Boca Raton, FL.
- Yi-Hsuan Yang and Homer H. Chen. 2012. Machine recognition of music emotion: A review. *ACM Transactions on Intelligent Systems and Technology* 3, 3 (2012), 40:1–40:30. DOI: <http://dx.doi.org/10.1145/2168752.2168754>

Received July 2016; accepted July 2016