D-JOGGER: SYNCING MUSIC WITH WALKING

Bart Moens
IPEM, Dept. of Musicology
Ghent University, Belgium
bmmoens.moens@ugent.be

Leon van Noorden
IPEM, Dept. of Musicology
Ghent University, Belgium
leonvannoorden@mac.com

Marc Leman
IPEM, Dept. of Musicology
Ghent University, Belgium
marc.leman@ugent.be

ABSTRACT

We present D-Jogger, a music interface that makes use of body movement to dynamically select music and adapt its tempo to the user's pace. D-Jogger consists of several independent modules, such as a step detection algorithm and tempo-aware playlists, to achieve this goal. The research done with D-Jogger has focused on entrainment: the synchronization of two rhythmical processes, in this case music and walking. We present several ways of visualizing entrainment data, including synchronization plots and phase histograms.

A pilot experiment was performed using D-Jogger with 33 participants. Preliminary data suggest that, when the music's tempo and the user's pace are close enough to each other, most users synchronize their walking to the music - taking a step with each beat. A user survey indicated that participants experience this effect as stimulating and motivating.

Several other application domains for D-Jogger are possible: personal training devices for joggers, rehabilitation therapy for Parkinson patients or simply as a nice-to-have application for your mobile phone.

1. INTRODUCTION AND RELATED WORK

Entrainment is broadly defined as a phenomenon in which two or more independent rhythmic processes synchronize with each other [2]. In a more constrained sense, it is the synchronization of a system with a variable frequency to an external frequency. We see this process every day, for example when people dance to music, they perform rhythmical movements in sync with the beat – on the perceived pulse of the music.

Another form of entrainment can happen when people are walking to music. Previous research has shown that people can synchronize their walking movements with music over a broad range of tempi and that this synchronization is most optimal around 120 beats per minute (BPM) [13]. Moreover, in [11] it is speculated that this frequency of 2 Hz represents some form of central resonance of human movement.

This paper describes the framework of D-Jogger. The framework is used to create applications that make use of the body movement of the user; specifically walking tempo (steps per minute, SPM). D-Jogger matches the tempo of the music (beats per minute, BPM) with the walking tempo of the user, switching songs when appropriate. The main hypothesis is that users will synchronize with the music by aligning their steps to the beats – a form of entrainment.

The idea of such a system is not novel in itself. Yamaha was the first to introduce BODiBeat[1] (2007), followed by Philips Activa[2] (2010). Both devices are focused on selecting music to optimize workout performance, depending on the user's pace and heart rate.

Several other systems have been proposed to do more or less the same thing: [7] and [5] use accelerometers to determine the user's pace, choosing a song from a preprocessed BPM-tagged music library. [7] has the ability to slow down or speed up the music without transposing the pitch. These papers focus on the technical aspect of the systems. [1] hints at the impact of such a system on the user, but the topic is not elaborate. However, none of these applications have been designed or used for research into entrainment, while this is the primary goal of D-Jogger.

2. D-JOGGER

D-Jogger is a framework employed to create a context-sensitive music player, using the user's pace to dynamically choose and adapt the music in real-time. The goal of D-Jogger is to be able to perform research into the phenomenon of entrainment; therefore the system must be flexible and easily adaptable.

Given this prerequisite, we opted to use Max/MSP in designing the application. Max/MSP is a graphical programming environment for real time audio processing that uses objects as basic building blocks, connectable via virtual wires. Objects developed for Max/MSP by third parties are called externals. D-Jogger consists of several externals, connectable in different ways to create a highly flexible framework for the rapid development of applications involving movement analysis and dynamic playlist generation. We give a short description of the available externals and their functioning; Figure 1 provides an overview.

---

[1] http://www.yamaha.com/bodibeat/

[2] http://www.consumer.philips.com/c/workout-monitors/176850/cat/us/

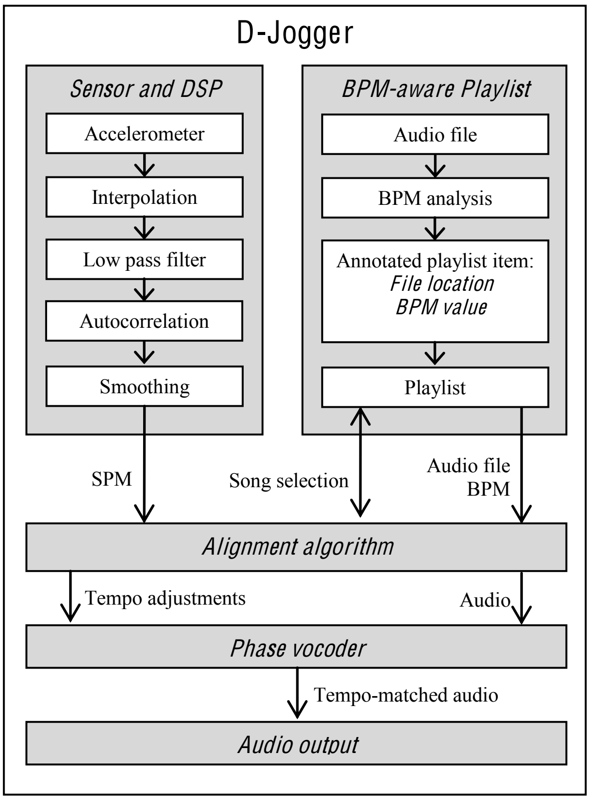


Figure 1. Overview of D-Jogger components

2.1. Sensor and Step Detection

A digital signal processing (DSP) external is used to extract the user's pace from a 3-axis ADXL330 accelerometer signal.

Accelerometer signals vary wildly depending on the gait pattern of the user, location of the sensor, the sensor specifications and the type of surface the user is walking on [8,9]. For example, soft grass, asphalt or treadmills result in different signals. Figure 2 illustrates the signal from different users walking on a treadmill with the sensor attached to their left ankle. The signal is from the axis perpendicular to the ground, showing clearly the moment of impact of the left heel on the treadmill. It is clear that both users' signals are periodic, but differ in other aspects. Several approaches are possible to determine step frequency, such as a Fourier transform or autocorrelation. A comparison of basic step detection algorithms for accelerometers can be found in [12]. For our purposes, we opted to use an autocorrelation algorithm.

The step detection algorithm consists of 4 phases:

- We interpolate the signal at a sample rate of 200 Hz with a second-order polynomial function;
- The resulting signal is filtered using a Finite Impulse Response (FIR) low pass filter of the 10th grade to reduce high frequency noise;
- We apply autocorrelation to the last four seconds of data, which contain at least 2 periods of the step signal. We then look for frequencies in the range of 30 to 125 SPM. Because the sensor is located on only one foot, the lower boundary of 30 SPM is equal to 60 SPM when taking both feet into account.
- The highest lag value is transformed to the SPM value. We take previous SPM values into account to reduce outlier effects, smoothing the results.

The resulting algorithm works independently from the selected axis of the sensor; however results are more stable if an axis with a clear periodic signal is selected.

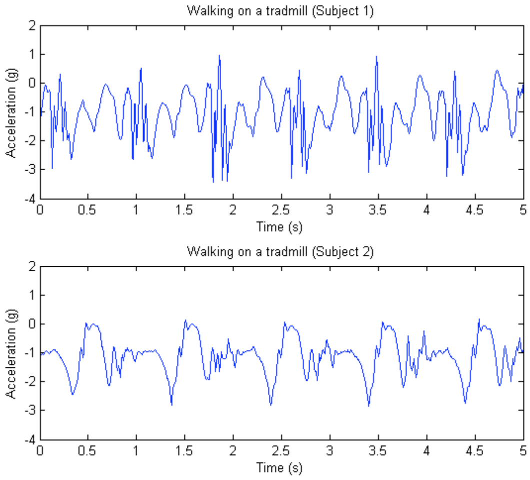


Figure 2. Acceleration signals of the left ankle of 2 subjects on a treadmill

2.2. BPM-Aware Playlist

The playlist external is used to dynamically add or remove songs from the playlist. When adding a song, it is first processed using BeatRoot [3] to extract the tempo information of the song. This includes timestamps of the beats and the BPM value. Optional metadata includes user ratings and general statistics on the usage of the song, preventing frequent repetitions of the same music.

2.3. Phase Vocoder

We use an external to time-stretch the audio without changing the pitch. Several solutions exist for this purpose [4]. ElasticX[3] is a Max/MSP object that allows changing the playback speed and pitch independently. Quality of the resulting audio is very good when the speed remains between 90% and 110% of the original. While this component is not part of our framework itself, it is used intensively and thus listed here.

2.4. Alignment algorithm

A central component connects all available externals. The alignment algorithm acts as the 'moderator' of the music: it decides whether the tempo of the current song should be adapted and whether a new song should be chosen. We present the Dynamic Song and Tempo (DSaT) algorithm, but many other alignment algorithms are possible.

---

[3] http://www.elasticmax.co.uk

DSaT consists of 2 phases:

- Song Selection: DSaT starts by choosing a song with a BPM close to the SPM value. If no suitable song is found, the SPM value is doubled and the search restarted. This feature is used with very low SPM values (SPM < 90), because very few songs with a BPM less than 90 are available [6].
- Dynamic Tempo: the required tempo adjustment for the song is calculated by dividing the SPM value by the BPM value. When this adjustment falls outside predetermined boundaries for more than 5 seconds, a new song is selected.

3. REPRESENTATIONS OF ENTRAINMENT

D-Jogger is used for research investigating human entrainment to music. We present several ways to visualize this entrainment, each having unique advantages and disadvantages. The following examples are visualizations of the exact timestamps of beats and footsteps from the pilot experiment. We first provide explanations of our chosen methods of visual representation. Results of the pilot experiment will be discussed in the next chapter.

3.1. BPM-SPM Plots

We create a plot where the BPM and SPM values are plotted over time. While this shows very little information about the entrainment itself, it is useful for validating the correct functioning of D-Jogger. We introduce the following notation: b_i is the timestamp of the i^{th} beat in milliseconds, s_j is the timestamp of the j^{th} step in milliseconds, bpm_i is the BPM value between b_{i-1} and b_i and spm_j is the SPM value between s_{j-1} and s_j (analogue for SPM calculation).

$$bpm_i = \frac{1000*60}{(b_i - b_{i-1})} \tag{1}$$

Figure 3 (top) gives an example of such a plot. Over a time span of 130 seconds, four different songs are played, as indicated by the red markers.

3.2. Synchronization plots

A synchronization plot shows the time between each step s_i and the closest beat b_j or b_{j+1}. We note b_j as the beat before s_i, b_{j+1} is the beat after s_i. The synchronization value, $synch_i$, is expressed in milliseconds: 0 means that the beat and the step occurred simultaneously; a positive value means the step occurred before the beat. A higher absolute value of $synch_i$ means a less optimal synchronization.

$$synch_i = \begin{cases} b_j - s_i & \text{if } \; s_i - b_j \le b_{j+1} - s_i \\ b_{j+1} - s_i & \text{otherwise} \end{cases} \tag{2}$$

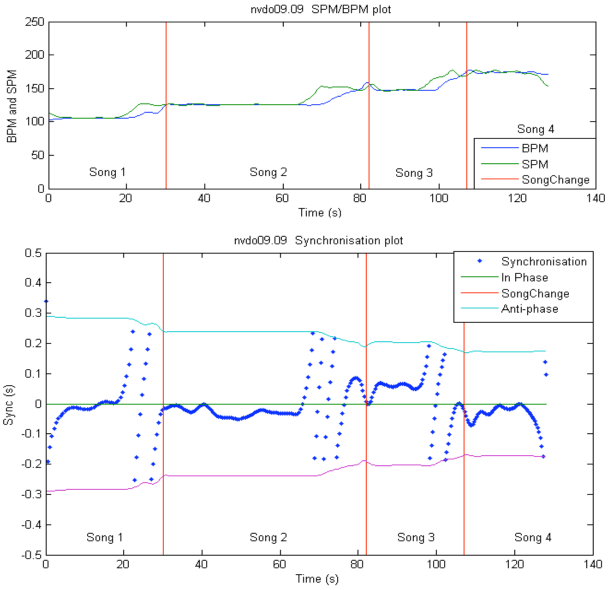


Figure 3. (top) BPM-SPM plot (bottom) Synchronization plot

When $synch_i$ equals or is close to 0, we speak of the footsteps and musical beats being *in phase*. When the steps occur exactly midway between two sequential beats, we speak of them being in *anti-phase.* The anti-phase time is inversely proportional to the BPM. For a visual representation, we plot anti-phase as well. The $synch_i$ value wraps between the positive and negative anti-phase.

Figure 3 (bottom) gives an example of a synchronization plot, representing the same data as the above BPM-SPM plot. This representation is very well suited to provide an overview of the users' actions. We can clearly see when the user is or is close to being in phase (e.g. between 5s and 20s in the graph). When the user speeds up it is clearly visible by the warping of the *synch* value (e.g. 20s to 30s). We can also see the inverse proportional relation of the anti-phase with the BPM value by comparing it with the top plot.

A disadvantage of the synchronization plots is that they are hard to compare to one another. This is due to the variables BPM and song changes. Therefore, we introduce a normalized representation called phase plots.

3.3. Phase plots

Based on a similar approach to that described in [10], we transform $synch_i$ to the phase angle θ_i by dividing by the BPM timing:

$$\theta_i = 2 * \pi * \frac{synch_i}{b_{j+1} - b_j} \tag{3}$$

In the visual representation, a circle is divided into 4 circular sectors. Details about the different sectors can be found in Table 1. The radius represents time, with $r = 1$ being the start of a song and $r = 0.3$ the end of a song. Time is spaced linearly between $r = 1$ and $r = 0.3$.

Using the same data as in the previous examples, we plot the third and forth song on a phase plot (Figure 4). In the figure we can see what we define as in phase and which steps fall within certain sectors. In the synchronization plot, walking appeared to be well synchronized to

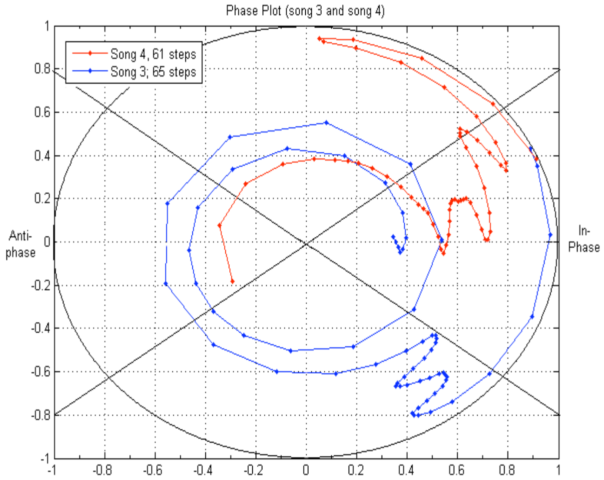
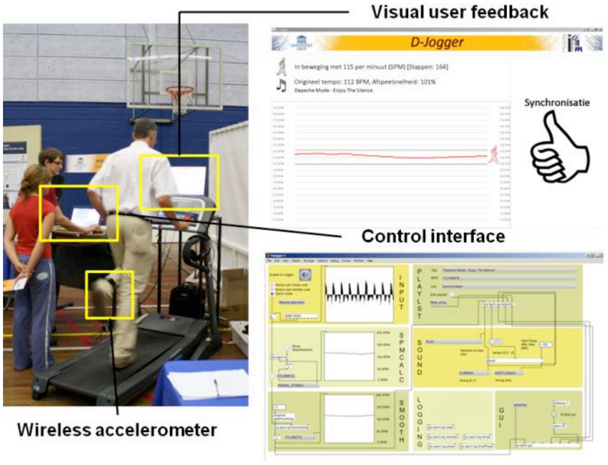


Figure 4. Phase plots of songs 3 and 4



In beweging met 115 per minuut (SPM) [Stappen: 164]
Origineel tempo: 112 BPM, Afspeelsnelheid: 101%
Depeche Mode - Enjoy The Silence

Figure 5. Pilot setup of D-Jogger

the beats of both songs. The phase plots reveal a different picture: the majority of the time, song 3 is not in phase, while the majority of song 4 is in phase. This is confirmed if we look at Table 1: the amount of time spent in a given sector is expressed in the last two columns. We can say that the user synchronized better in the fourth song than in the third. Phase plots and tables make it easier to compare different sets of synchronization data both visually and numerically because of the BPM normalization.

Circular sector	Angles	Song 3	Song 4
In-Phase	0°:39° & 320°:359°	20.31%	66.67%
Anti-phase	140°:219°	7.81%	3.33%
Rest	40°:139° & 220°:319°	71.86%	30.00%

Table 1. Definition of circular sectors on phase plots and two examples of time spend in each sector for songs 3 and 4

3.4. Phase histogram plots

Next, we create a histogram of the phase angles. A phase histogram gives a global overview of the amount of time a user spends in a certain phase angle, giving an overview of the user's entrainment. Figure 6 illustrates the phase histogram for all available data from the pilot experiment. The advantages of this representation are clear: independent of the number of steps and easily comparable with different histograms, they give a clear overview of entrainment. These results can also be represented in a table for easier statistical analysis.

4. PILOT EXPERIMENT

4.1. Setup

We designed a version of D-Jogger for demonstration purposes, using a 3-axis ADXL330 accelerometer attached at the user's ankle and connected via Bluetooth. A graphical user interface (GUI) was placed in front of the user, showing the user's SPM history, music information and the synchronization of BPM and SPM. The GUI did not include information about the entrainment itself. Music was chosen from a library containing 50 pop songs. The DSaT algorithm was used for music selection and real-time adaptation. If the necessary tempo adjustment was outside the boundaries of ±10% for more than 5 seconds, a new song was chosen with a BPM closer to the current SPM. These boundaries were also shown in the GUI. After using D-Jogger, participants were asked to complete a survey. Figure 5 illustrates the experimental setup.

4.2. Participants, experiment and survey

A total of 33 participants experimented with the system at a public event in 2009. The participants were given a short introduction prior to the experiment, explaining the concept of time-stretching music to the users' pace. Participants were not informed of the concept entrainment.

After the introduction, the sensor was attached to their left ankle. Once on the treadmill, users were free to experiment with D-Jogger, picking and changing their speed without time limitations. The treadmill control board featured speed presets (2, 4 … 20 km/h) and fine speed control (0.1 km/h).

After the experiment, participants filled in a survey asking their personal jogging preferences and their experience with D-Jogger.

4.3. Data and analysis

A Zoom H4 stereo recorder was used to record the whole experiment. One channel recorded the audio as perceived by the user; the other channel recorded the footsteps using a microphone placed near the front of the treadmill. This was done to have a dataset independent of the D-Jogger logs in order to verify the correct functioning of D-Jogger itself. From the recordings we manually extracted the timestamps of beats and steps. A total of 15 datasets were usable, giving a total of 3700 steps. Various parts of the recording were unusable because of unclear step recordings.

4.4. Results

We calculated the synchronization phase for 3700 steps. Figure 6 depicts the results in a phase histogram. This shows that the majority of steps (56.79 %) were in sync with the music. Because D-Jogger does not do phase synchronization, the user was responsible for the entrainment effect between the music and their gait pattern. 17.37% steps were taken in anti-phase, the remaining 25.85% of steps were neither in phase or in anti-phase and mostly occurred close to song changes.

Our preliminary conclusions are that:

- While using D-Jogger, the majority of subjects synchronized to the beats, no matter what the users' pace.
- If the music tempo is close enough to the user's pace, the user tends to synchronize his or her steps with the beats.
- Users reported in the survey that they feel more motivated when in sync with the music.

5. DISCUSSION

The question arises as to whether a treadmill stimulates or hinders entrainment. One might say that the SPM of the user is dependent on the speed selected on the treadmill. While this is partially true, a user can speed up or slow down the SPM value significantly by adjusting the step length, as seen in several datasets. This enables the user to synchronize with the music on a treadmill. However, we do not know to what extent the use of a treadmill influences the entrainment process. Future studies will include a comparison of uninhibited locomotion activity and locomotion activity on a treadmill.

While this was as small study, some interesting results were procured. They do require confirmation by new experiments in more rigorous settings, but the pilot experiment gives us hints about the expected results. The result of the survey, where users reported feeling more motivated when in sync with the music, is compelling but also subjective and requires further validation.

Currently, the prototype of D-Jogger has only been tested using songs with a typical 4/4 time signature and constant beat. This is partly due to the limitations of the beat extraction, which works optimal with the simple and more popular Western songs. With the goal of D-Jogger in mind, we require songs with an unambiguous and clear rhythm, so in this early stage we do not see this as a problem.

Another important remark is that when a new song is chosen, phase synchronization could be temporarily lost. Currently, a new song starts independently of the time when footsteps occur, so there is a high chance that the user will be out of sync with the music immediately following a song change. For this we intend to analyze the sensors for phase information of the steps so a song can be started in accordance with the step phase.

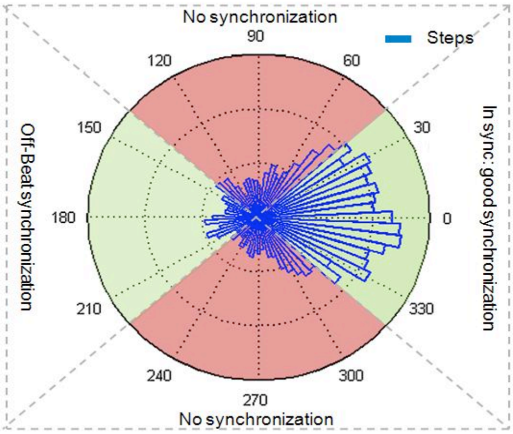


Figure 6. Results of the pilot experiment

6. FUTURE WORK

The most limiting factor up until now has been the use of a stationary version of D-Jogger with a treadmill. A mobile version will be implemented, using the internal accelerometer of the mobile device. With the mobile version we can perform an experiment with unconstrained locomotion, yielding results closer to our typical real-life experience with walking to music.

The mobile version will also feature a true multimodal interface: user control will be possible using simple gestures on a touch screen. This multimodal interface is ideal for use in training sessions because the user can select and rate music without interrupting the training.

The D-Jogger framework can also be used in other applications. We are currently looking into rehabilitation therapy for Parkinson patients using audio cues and a personal training device.

7. CONCLUSIONS

While there is still much to be done, D-Jogger shows great potential in both the current application and future possibilities involving entrainment research in body movement in relation to music.

We proposed several options to visualize entrainment data. BPM-SPM plots show the evolution of tempo of both the music and the user in time, while the complementary synchronization plots show the evolution of entrainment in time. Both are very useful for visual inspection, while phase tables, plots and histograms are better suited for analytical purposes. They show the synchronization data in a normalized way, allowing comparison of different datasets.

We performed a pilot study using the D-Jogger. The results show our hypothesis, that entrainment occurs when SPM is close to BPM, to be correct. However, further research is necessary using both the stationary and mobile version of D-Jogger to confirm the hypothesis.

REFERENCES

[1] J.T. Biehl, P.D. Adamczyk and B.P. Bailey, “DJogger: a mobile dynamic music device”. *In CHI '06 Extended Abstracts on Human Factors in Computing Systems* pp. 556-561, 2006

[2] M. Clayton, R. Sager and U. Will, “In Time with the Music: The Concept of Entrainment and its Significance for Ethnomusicology”, *European Meetings in Ethnomusicology*, Vol. 11, pp. 3–142, 2005.

[3] S. Dixon, “Evaluation of the Audio Beat Tracking System BeatRoot”, *Journal of New Music Research*, Vol. 36, pp. 39-50, 2007

[4] D. Dorran, B. Lawlor and E. Coyle, “A comparison of time-domain timescale modification algorithms”, *In 120th Audio Engineering Society Convention*, issue paper 6674, 2006.

[5] G.T. Elliott, B. Tomlinson, “PersonalSoundtrack: context-aware playlists that adapt to user pace”, *CHI '06 extended abstracts on Human factors in computing systems*, 2006.

[6] F. Gouyon, “Dance music classification: A tempo-based approach”, *In Proceedings of the International Conference on Music Information Retrieval*, Vol. 2004, pp. 501-504, 2004.

[7] J. Hockman, M.M. Wanderley and I. Fujinaga, “Phase vocoder manipulation by runner’s pace”, *In Proceedings of the 2009 Conference on New Instruments for Musical Expression*, pp. 90-93, 2009.

[8] L. Hongyin and T. Kaiyu, “The reliability of using accelerometer and gyroscope for gait event identification on persons with dropped foot”, *Gait & Posture*, Vol. 27, Issue 2, pp. 248-257, 2008.

[9] H.-J Jang, J.W. Kim and D.-H. Hwang, “Robust step detection method for pedestrian navigation systems”, *Electronics Letters*, vol. 43, no. 14, 2007.

[10] J.A.S. Kelso, *Dynamic Patterns: The Self-Organization of Brain and Behavior*. Cambridge, MA: The MIT Press, 1995.

[11] H. G. MacDougall and S. T. Moore, “Marching to the beat of the same drummer: the spontaneous tempo of human locomotion”, *Journal of Applied Physiology,* vol. 99, pp. 1164-1173, 2005.

[12] M. Marschollek, M. Goevercin, K.H. Wolf, B. Song, M. Gietzelt, R. Haux and E. Steinhagen-Thiessen, “A performance comparison of accelerometry-based step detection algorithms on a large, non-laboratory sample of healthy and mobility-impaired persons”. *Conference Proceedings IEEE Eng Med Biol Soc 2008*, pp. 1319-1322, 2008.

[13] F. Styns, L.V. Noorden, D. Moelants and M. Leman, “Walking on music”, *in Human Movement Science*, Vol. 26, Issue 5, pp. 769-785, 2007.