Multimodal Detection of Music Performances for Intelligent Emotion Based Lighting

Esben Oxholm, Ellen K. Hansen, and Georgios Triantafyllidis(✉)

Lighting Design - Architecture, Design and Media Technology, Aalborg University Copenhagen, AC Meyers 15, 2450 Copenhagen, Denmark
ebonde14@student.aau.dk, {ekh,gt}@create.aau.dk

Abstract. Playing music is about conveying emotions and the lighting at a concert can help do that. However, new and unknown bands that play at smaller venues and bands that don't have the budget to hire a dedicated light technician have to miss out on lighting that will help them to convey the emotions of what they play. In this paper it is investigated whether it is possible or not to develop an intelligent system that through a multimodal input detects the intended emotions of the played music and in real-time adjusts the lighting accordingly. A concept for such an intelligent lighting system is developed and described. Through existing research on music and emotion, as well as on musicians' body movements related to the emotion they want to convey, a row of cues is defined. This includes amount, speed, fluency and regularity for the visual and level, tempo, articulation and timbre for the auditory. Using a microphone and a Kinect camera to detect such cues, the system is able to detect the intended emotion of what is being played. Specific lighting designs are then developed to support the specific emotions and the system is able to change between and alter the lighting design based on the incoming cues. The results suggest that the intelligent emotion-based lighting system has an advantage over a just beat synced lighting and it is concluded that there is reason to explore this idea further.

Keywords: Multimodal detection · Emotion-based lighting

1 Introduction: Music, Emotion and Light

The origin, meaning and purpose of music is not clear cut. Though, one common theme seems to be agreed upon: Music convey emotions. As Juslin and Laukka [1] puts it: *"[…] A convincing emotional expression is often desired, or even expected, from actors and musicians"*.

Music is capable of generating and amplifying the feelings of being happy, sad, angry, motivated, etc. [2]. Examples of how music is used in our daily lives underscores this statement. Think of the nature of the music used at e.g. a funeral versus a summer party, or how the music in a cozy café seeks to generate a relaxed and laid back atmosphere. Another example is the way music effectively is used in movies to enhance the intended emotion of a particular incident.

A.L. Brooks and E. Brooks (Eds.): ArtsIT/DLI 2016, LNICST 196, pp. 212–219, 2017.
DOI: 10.1007/978-3-319-55834-9_25

If one of the main purposes of music is to convey emotions, then it must be assumed that the purpose of live concerts is to convey emotions as well. When listening to a record, only the auditory sense is being used as a channel for receiving the emotional output. At live concert the visual sense is brought into play as well.

Several visual tools can be used to express certain emotions. One is, that apart from playing the music, the performers are able to use their body language and facial expressions to convey emotions. Another is the outfit of the performers and the stage decoration. It can be done in a way that brings out certain emotions and moods. A third factor is lighting. Whether intended or not the lighting has an effect on the conveyed emotions.

At the time being, research that either approves or disproves this third statement has not been found. Therefore, it is treated as a hypothesis. However, there seems to be a common agreement upon that lighting is important for the emotions communicated at concerts:

> *"Lighting at a rock show is one of those things that most people don't consciously think about but can strongly impact their experience. Through lighting techniques, a stadium show can feel as intimate as a small club show. The audience can be made to feel inspired, disturbed, moved, or impressed, depending on the emotions that the artist is trying to communicate. It is one of the many sophisticated ways that the rock concert industry has developed to enhance the live concert experience and make it truly spectacular"* [3].
>
> *"Lighting adds as much to the experience as the artist we have paid money to see. Nowadays, when we perhaps spend much more of our money on seeing our favorite musicians live than we do on expanding our CD collection, the role of lighting in these events has never been more under the spotlight"* [4].

A concert is always accompanied by lighting in someway. From advanced pre-programmed light shows at big stadium concerts with an audience of tens of thousands to a few static light sources at small stage in a café with maybe five people (not) looking. Even when there is no dedicated lighting for the concert, there will always be the daylight or the artificial light that is in the room where the music is performed. Whether the lighting helps to convey the intended emotions or not, depends on if the lighting fits with the intended emotions of the music. E.g. if the purpose of the music being performed on a small stage in a café is to provide relaxing background music, then a few static bulbs helps to enhance the feeling of relaxation. No movements, no blinking lights and no shifting colors. In the same way, if the purpose of the stadium rock concert is to convey emotions of excitement, control and anger, then the advanced pre-programmed light show seems like an optimal way to enhance that. A lot of variation, shifting colors, blinking lights. Try to imagine the opposite. A wild light show for the relaxing café concert and a few static bulbs for the rock concert. It would not fit well with the music.

This statement is supported by Ethan Weber, who has designed lighting for Marilyn Manson, My Chemical Romance and Green Day and operated lights for The Rolling Stones and U2. When talking about one of his successful lighting design he ends concluding: *"The lighting worked because it matched the emotional meaning of the song"* [5].

In this context, this paper tries to answer the question of how to create an intelligent lighting system that helps conveying the intended emotion of what is being played at a concert and thus enhancing how well the emotions are being perceived by the audience. Answering this question will provide the added value of enabling bands on a budget

playing at small venues to give a stronger emotional live performance through lighting. At the time being, this work does not seek to compete with a dedicated light technician or a fully programmed light show. Rather it should be seen as an improvement to the beat-synced lighting and ultimately as an option to the in-house light technician.

The rest of the paper is organized as follows: The next section analyses the emotion cues (both auditory and visual) that are perceived by the audience. Then, the proposed design solution for the lighting control is presented in Sect. 4 to support conveying the intended emotion at a concert. Next section presents a proof of concept test and finally conclusions are discussed in the last section.

2 Analysis: Emotion Cues

When an audience are looking at a performer or a band playing a song, the two most activated senses are the auditory and the visual.

Auditory emotion cues: Research on music and emotion shows that there exists several cues in music that can be used to determine the emotion of what is being played. As Juslin and Sloboda write [6]: *"[…] researchers have tried to describe the means by which performers express specific emotions […] One main finding from this line of research is that the performer's expressive intention affects almost every aspect of the performance; that is, emotional expression in performances seems to involve a whole set of cues - or bits of information - that are used by performers and listeners."*

Also, Juslin and Sloboda have summed up the results in a diagram showing the auditory cues used to express happiness, sadness, anger, fear and love/tenderness. These are the emotions that have been studied most and furthermore are regarded as 'basic-emotions' by scientists [6]. As shown in Fig. 1, a vast amount of different cues affects the intended emotion. However, they are not equally important and they are not always present. They end up concluding that the cues that have the greatest impact on the

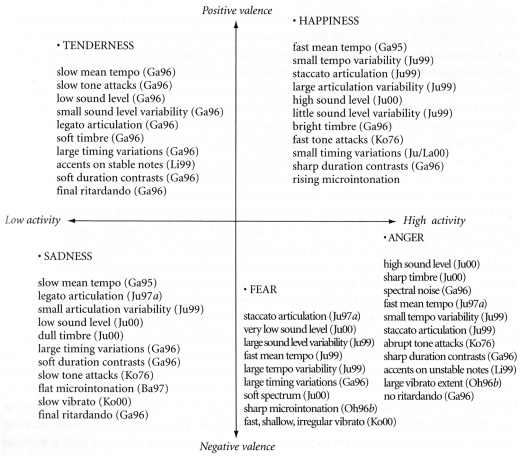

Fig. 1. Auditory cues

intended emotions is tempo, sound level, articulation and timbre. It is thus possible to define the emotion of a piece of music by analyzing these four cues.

Visual emotion cues: In the same way as the detection of emotion from sound, the detection of emotion from a musician's body movement relates to a set of cues. In a study from 2007, Dahl and Friberg [7] investigate the visual information provided by the performer. As described in the study paper: *"Musicians often make gestures and move their bodies expressing the musical intention. This visual information provides a channel of communication to the listener of its own, separated from the auditory signal. In order to explore to what extent emotional intentions can be conveyed through musicians' movements, subjects watched and rated silent video clips of musicians performing four different emotional intentions, Happy, Sad, Angry, and Fearful"*.

Besides rating how well the emotions were conveyed, the subjects were also asked to rate the movement they saw, based on the following four cues: Regularity, fluency, speed, amount: "*The assumption was that Amount would correspond to an overall measure of the physical magnitude of the movement patterns, Speed to the overall number of movement patterns per time unit, Fluency to the smoothness of movement patterns, and Regularity to the variation in movement patterns over the performance*". The study in [7] ends by concluding that the emotions of sadness, happiness and anger was successfully conveyed through body movements only, while fear was not. It is also concluded that the cues mentioned above can be used to describe the different intended emotions.

Overview of the emotion cues: The auditory and visual cues used to convey the emotions of happiness, sadness and fear are summed up in Table 1.

Table 1. Auditory and visual cues

		Happiness	Sadness	Anger	Fear
Auditory cues	*Tempo*	Fast	Slow	Fast	Fast
	Sound level	High	Low	High	Very low
	Articulation	Staccato	Legato	Staccato	Staccato
	Timbre	Bright	Dull	Sharp	n/a
Visual cues	*Regularity*	Regular	Regular	Regular	n/a
	Fluency	In-between jerky and smooth	Smooth	Jerky	n/a
	Speed	Fast	Slow	Fast	n/a
	Amount	Medium	Small	Medium	n/a

3 Proposed Design

Having established the most important cues that are used to convey certain emotions, it is time to focus on how the intelligent system should detect and process these cues (which for sound is tempo, sound level, articulation and timbre and for body movement is regularity, fluency, speed and amount). Basically the system can be divided into three stages: Input, processing and output (see Fig. 2):

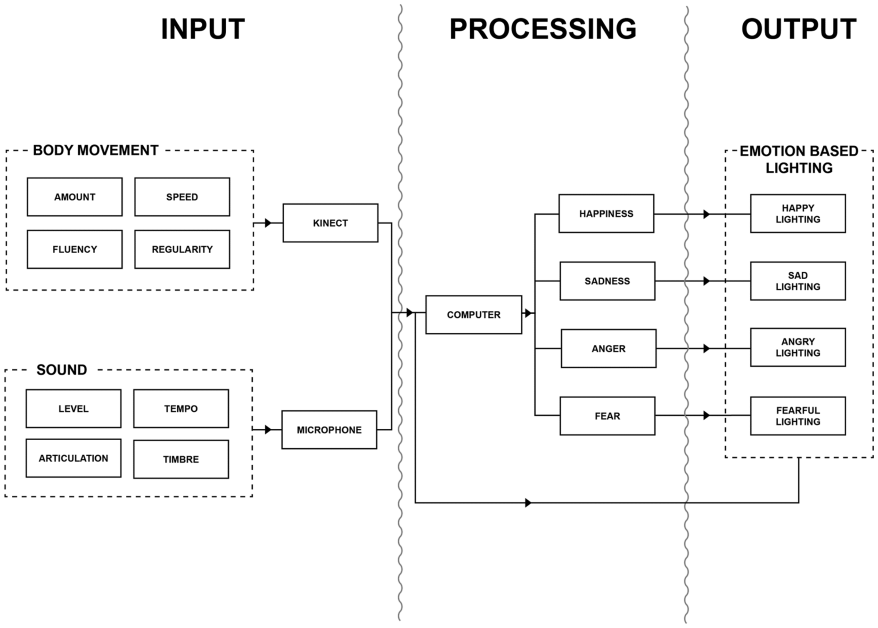

Fig. 2. System's flowchart

Input: The input stage deals with the cues described above. The visual cues are going to be detected via a Microsoft Xbox 360 Kinect camera (using the Processing library SimpleOpenNI). The advantages of using the Kinect over a regular camera is the use of infrared light. The sound is detected through a regular microphone.

Processing: In the stage of processing, the detected cues are fed into a piece of software where they are analyzed. Based on predefined values in the software, the system is able to decide if the cues together result in an emotion of happiness, sadness, anger or fear.

Output: The third stage of the system is the output of the lighting. The software contains different lighting designs that are developed specifically to the different detected emotions. E.g. if the emotion detected based on the cues is anger, then a lighting design that supports the emotion of anger is activated. The lighting design for each emotion is not static. The cues detected in the input stage are used to alter the lighting within certain boundaries for each emotion.

4 Proof of Concept Test

For the design testing, a simplified proof-of-concept version of the system is being used. That is to keep the number of variables low and thus keeping the number of possible errors as low as possible. If it is possible to make a system that works using only a couple of cues, then the system can be expanded to include all cues and emotions. To this goal, the possible emotions has been reduced to happiness and sadness. The reason for this

choice is the fact that they are the ones that consists of cues that are most different. Also, it seems to be the emotions that are the most basic to express and recognize [7]. In the input stage the cues have been reduced to the speed of the body and, level and tempo of the music. The speed of the body has been chosen because that it is the visual cue that differ the most between sadness and happiness. The level and tempo of the music have been chosen as they seem to be the most common and easiest recognizable cues.

So, how is lighting designed to support emotions of happiness and sadness respectively? As described earlier, the emotion in music can be described via sets of certain cues coming from both auditory and visual input. A set of cues can be defined for the lighting as well, that in the same way as with music, defines the intended emotion. Research has shown that the cues used for speech relates to the cues used in music: *"The results revealed a number of similarities in code usage. For example, vocal expression of sadness is associated with slow speech rate, low voice intensity, low intonation, and little high-frequency energy in the spectrum of the voice."* [6]. Also, a research study on dance, or people's movement to music, concludes that happy movements relates to high-dimensional movements, while sad movements is simple, low dimensional, long and smooth movements and covering little space [8]. Based on the mentioned research and the recurrence of almost identical emotion cues across disciplines, it is hypothesized that designing the lighting based on the same cues will make sure that it supports the intended emotions of the performance.

Light attributes: The cues used to define the emotion of the lighting are inspired by the most important auditory and visual cues, which is *level, tempo, articulation and timbre* and *amount, speed, fluency and regularity*. The ones used for the lighting are defined as: *Intensity, speed, fluency, regularity, hue, saturation and brightness.* The intensity relates to the brightness of the output. The speed defines how many changes, within a certain emotion, happens over a period of time. The fluency relates to how fluent the changes between different intensities and hues are within a certain emotion. The regularity relates to variation in patterns over the performance. Hue and saturation relates to the hue and saturation of the color of the lighting.

Happy lighting: The cues from previous research that relates to happy is defined as fast tempo, high level, staccato articulation, bright timbre, regular, somewhat jerky fluency and medium amount of movement. These cues have been used to define the cues for happy lighting shown in Table 2.

Table 2. Lighting cues

Emotion	Intensity	Speed	Fluency	Regularity	Hue	Saturation
Happiness	High	Fast	Somewhat jerky	Regular	Yellow	High
Sadness	Low	Slow	Smooth	Regular	Violet	Low

Sad lighting: As with the happy lighting, the cues for sad lighting has been defined inspired by the cues of the previous research. That is, slow tempo, low level, legato articulation, dull timbre, smooth fluency and small amount of movement. The cues defined for the sad lighting can be seen in Table 2.

The hue of the lighting has been inspired by a research study on how colors relates to music conveying different emotions [9]. Although the results are not very strong, the study concludes that the hue that relates the most to happiness is yellow and the one that relates most to sadness is violet.

Test: For comparison reasons the emotion based lighting will be tested against beat-synced lighting. As written in the introduction, one of the goals for the emotion-based lighting is to be an improvement to beat-synced lighting. To this goal, prior to the test the performer prepared two solo performances for a single snare drum. The two performances are exactly the same composition-wise and rhythmically. The difference is the fact that the first rhythm is played using the cues that relates to sadness and the second rhythm is played using the cues that relates to happiness. I.e. rhythm 1 is slow, low in volume and requires slow body movements, while rhythm 2 is fast, loud and uses fast body movements.

Findings: Figure 3 shows the results from the test. The orange dots represent the markings from audience 1 who were subject to the emotion-based lighting. The grey dots represent audience 2 who were subject to the beat-synced lighting. The top line shows the results for the first performed rhythm, the sad one, while the bottom line shows the performance of rhythm 2, the happy one.

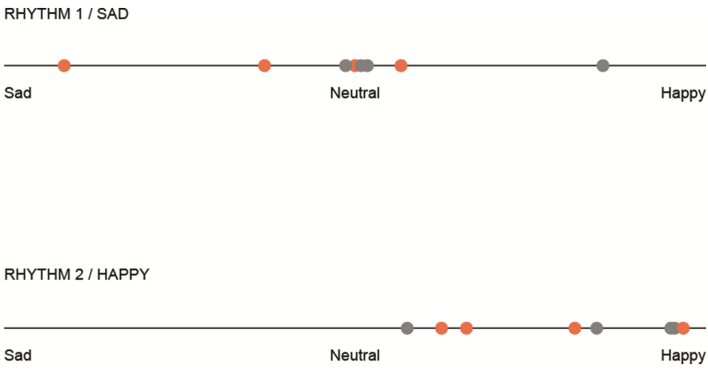

Fig. 3. Test results (Color figure online)

Although this test is based on a small sample size, it supports the idea that the emotion-based lighting in fact are better at conveying the intended emotion of what is being played and thus enhancing how well the emotions are being received by the audience compared to beat-synced lighting. The results show that the beat-synced lighting failed in conveying the emotion of sadness. One test subject even perceived the sad performance as being happy. However, the beat-synced lighting was as good as the emotion-based lighting at conveying happiness.

5 Conclusions

Through existing research on music and emotion and musicians body movements related to the emotion the want to convey, a set of cues to look and listen for was defined. This included amount, speed, fluency, and regularity for the visual and level, tempo, articulation, and timbre for the auditory. Using a microphone and a Kinect camera to detect these cues it was possible to create a system that is able to detect the intended emotion of what is being played - at least in a proof-of-concept version dealing with the emotions of happiness and sadness.

Specific lighting designs were developed to support the specific emotions and the system was able to change between and alter the lighting based on the incoming cues. It was tested how well the emotion-based lighting performed in enhancing the conveying of the intended emotions compared to beat-synced lighting. The results showed that the emotion-based lighting is promising.

References

1. Juslin, P., Laukka, P.: Communication of emotions in vocal expression and music performance: different channels, same code? Psychol. Bull. **129**(5), 770–814 (2003)
2. Trainor, L.J., Schmidt, L.A.: Processing emotions induced by music. In: The Cognitive Neuroscience of Music, pp. 310–324 (2003)
3. Lighting the Rock Concert Corporate Event. http://www.forbes.com/sites/ruthblatt/2014/08/25/lighting-the-rock-concert-corporate-event/#1dc48e2bf90e
4. Playing with the Spectacular Interactive Lighting. http://origin.www.futureoflight.philips.com/post/84508677106/playing-with-the-spectacular-interactive-lighting
5. Lighting the Rock Concert Corporate Event. http://www.forbes.com/sites/ruthblatt/2014/08/25/lighting-the-rock-concert-corporate-event/#2715e4857a0b6c165f59f90e
6. Juslin, P., Sloboda, J.: Music and Emotion, Theory and Research. Oxford Press, New York (2001)
7. Dahl, S., Friberg, A.: Visual perception of expressiveness in musicians' body movements. Music Percept. **24**(5), 433–454 (2007)
8. Burger, B., Saarikallio, S., Luck, G., Thompson, M.R., Toiviainen, P.: Emotions Move Us: Basic Emotions in Music Influence People's Movement to Music (2012)
9. Bresin, R.: Real-time visualization of musical expression. In: Proceedings of HUMAINE Workshop on From Signals to Signs of Emotion and Vice Versa', Santorini, Greece. Institute of Communication and Computer Systems, NTUAthens (2004)