Chapter 2
Representations of Emotions

2.1 Perception of Emotions

Emotions are a dominant element in music, and they are the reason people listen to music so often. We can ask ourselves the question: What can be the possible perception of emotions while listening to music? The psychologist Gabrielsson in his work [25] made a distinction between emotion perception into perceived and felt (induced) emotions. In the case of the former, we can perceive emotional expression in music without necessarily being affected ourselves; while in the latter, we have an actual emotional response to the music. Perceived emotion is the emotion recognized in the music, and induced emotion is the emotion experienced by the listener. Perceived and felt emotions are two alternatives that were the focus of psychology papers, such as those by Juslin and Laukka [48] and by Vuoskoski [108].

In our own work analyzing music recordings, we consider perceived emotion in music. During our experiments, experts with a university music education were asked to describe the emotions they perceived in music fragments and their opinions were then used to build a model of emotion prediction in music recordings.

2.2 Categorical Approach

Music emotion detection studies are mainly based on two popular approaches: categorical or dimensional. In the first, emotions are described with a discrete number of classes, affective adjectives, and in the second emotions are identified by axes. In the categorical approach, there are many concepts about class quantity and grouping methods. One of the first psychology papers that focused on finding and grouping terms pertaining to emotions was by Hevner [42]. As a result of the conducted experiment, there was a list of 66 adjectives arranged into eight groups distributed on a circle (Fig. 2.1). Adjectives inside a group are close to each other, the nature of
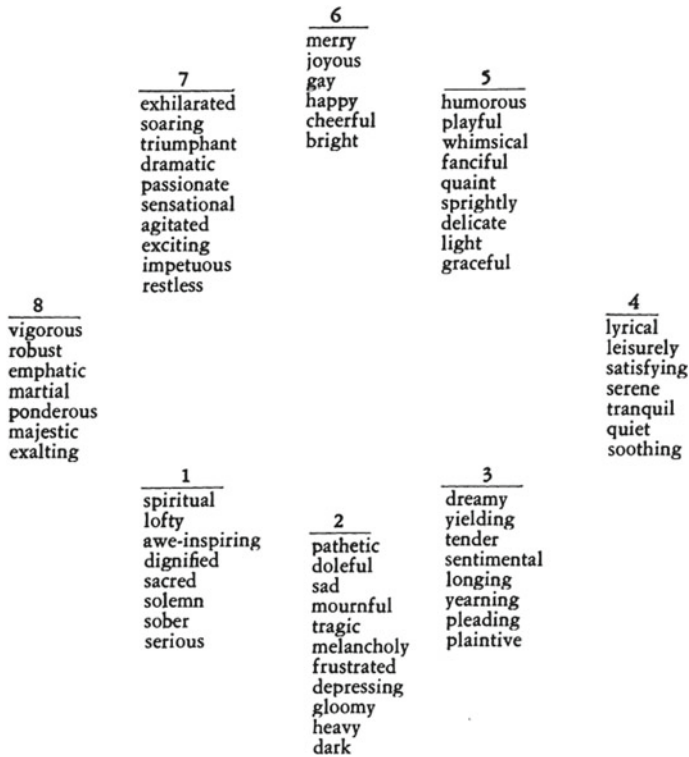
Fig. 2.1 Hevner's adjectives arranged in eight groups [42]

adjacent groups is evolving, and opposite groups on the circle are the furthest apart by emotion. Hevner's model was then modified by Farnsworth [23] and Schubert [97], who decreased the number of adjectives to 50 and 46, grouped them into nine groups.

Another interesting and important concept of finding the category of an emotion is the concept of basic emotion presented by Ekman [21, 22], which was developed for facial expression. Ekman describes features that enable differentiating basic emotions, which are:

- happiness,
- sadness,
- anger,
- fear,
- disgust,
- and surprise.

Ekman conducted experiments proving that facial expressions of basic emotions are cross-cultural. Johnson-Laird and Oatley [47] presented a somewhat smaller group of basic emotions: happiness, sadness, anger, fear, and disgust.

In the community of Music Information Retrieval Evaluation eXchange (MIREX) for automatic music mood classification, five mood clusters were used for song categorization [43]:

- Cluster 1 (passionate, rousing, confident, boisterous, rowdy);
- Cluster 2 (rollicking, cheerful, fun, sweet, amiable/good natured);
- Cluster 3 (literate, poignant, wistful, bittersweet, autumnal, brooding);
- Cluster 4 (humorous, silly, campy, quirky, whimsical, witty, wry);
- Cluster 5 (aggressive, fiery, tense/anxious, intense, volatile, visceral).

Hu et al. in [44] indicates, however, that the clusters might not be optimal and noticed some semantic overlap; similar findings were noted by Chen et al. [14]. The research carried out by Laurier et al. [55, 56] indicates deficiencies in this categorization, for example: experiments found that Cluster 1 and Cluster 5 are quite similar.

A popular emotion set used to categorize emotions in music turned out to be a collection consisting of 4 classes: happy, angry, sad, and relaxed. It corresponds to the four quarters of Russell's model [88], which were formed by dividing a plane by two perpendicular axes: arousal and valence. These values clearly define a point on the plane corresponding to a specific emotion and locate it on one of four quarters of Russell's model. The basic classes of emotions are assigned to the quarters as follows:

- happy—arousal high, valence high;
- angry—arousal high, valence low;
- sad—arousal low, valence low;
- relaxed—arousal low, valence high.

The selection of four categories of emotions also refers to the theory of basic emotions presented by Ekman [21]. The four categories are representatives of the main emotions from each of the quarters.

A significant disadvantage of the categorical approach is that the number of emotions and their shades perceived in music is much richer than the limited number of categories of emotions. The categorical approach has poorer resolution, by using the categories, we simplify the description of emotions in music, which facilitates understanding the character of the emotions and provides only a general overview of the emotions in music. One category contains an entire set of various shades of emotions. The smaller the number of groups in the categorical approach, the greater the simplification.

In this work, a set of four basic emotions: happy, angry, sad and relaxed, corresponding to the four quarters of Russell's model, were used for the analysis of music recordings using the categorical approach.

2.3 Dimensional Approach

In the dimensional approach, emotions are identified on the basis of their location in a space with a small number of emotional dimensions. In this way, the emotion of a song is represented as a point on an emotion space.

The two-dimensional circumplex model of emotion, which uses the two dimensions of arousal and valence, was presented by Russell in [88]. Arousal could be high or low and valence positive or negative (Fig. 2.2). In this model, all emotions can be understood as changing values of valence and arousal.

A variant of Russell's model is Thayer's model [103], in which the author suggested that two basic dimensions of describing emotions are two separate arousal dimensions: energetic arousal and tense arousal. In Thayer's model, valence could be explained as varying combinations of energetic arousal and tense arousal. Figure 2.3 is a visual presentation of the two models.

An example of a model where an emotion is described using three dimensions is Mehrabian and Russell's Pleasure-Arousal-Dominance (PAD) model [67], which was originally constructed to measure a person's emotional reaction to the environment. The three basic dimensions of emotions and their descriptions are: pleasure—positive and negative affective states; arousal—energy and stimulation level; dominance—a sense of control or freedom to act.
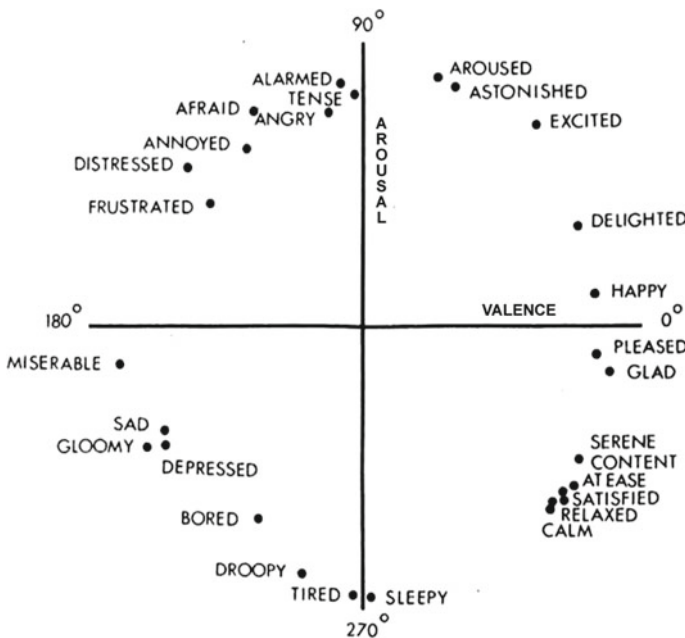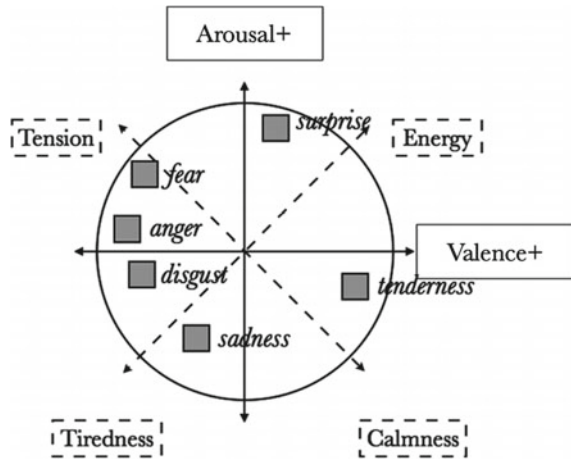

Fig. 2.2 Russell's circumplex model [88]

Fig. 2.3 Dimensional models of emotions with common basic emotion categories overlaid. In Russell's model, the axes are indicated by a solid line; in Thayer's model, the axes are indicated by a dotted line [20]

Experiments on automatic emotion prediction in music files using the PAD (Pleasure-Arousal-Dominance) and PA (Pleasure-Arousal) models were conducted by MacDorman et al. [64]. During file indexing, the authors noticed a significant correlation between the values of arousal and dominance. Ultimately, they decided to abandon the dominance dimension and use the Pleasure-Arousal model, because the results indicate that the dominance dimension was not informative for music.

A comparison of the discrete and dimensional models of emotion in music is presented by Eerola and Vuoskoski in [20], who used in their experiments five discrete emotions (anger, fear, sadness, happiness, and tenderness) and three bipolar dimensions (valence, energy arousal, and tension arousal). Linear mapping techniques between the discrete and dimensional models revealed a high correspondence along two central dimensions (valence and arousal). They concluded that the three dimensions could be reduced to two without significantly reducing the goodness of fit.

2.4 Summary

In our own work analyzing music recordings, we consider perceived emotion in music. Two approaches, categorical and dimensional, were used in emotion detection experiments. Using the categorical approach, a set of four basic emotions—happy, angry, sad and relaxed—were used. Using the dimensional approach, Russell's model—the most universal and least complicated to apply—was used. The four quarters of Russell's model correspond to four categories of emotions used in the categorical approach, which combines these two approaches to a certain degree. The categorical approach is more general and simplified in describing emotions, and the dimensional approach is more detailed and able to detect shades of emotions.