

Émotions dans la musique : approche basée sur le contenu

Massinissa Hamidi, Hassane Gaci, Van Luan Nguyen

Université Paris Diderot - Paris 7

07 Décembre 2016

Motivations

- La musique est un énorme vecteur d'émotions
- Curiosité, explorer d'autres types de recommandation moins conventionnels
- intérêt grandissant pour les émotions de la part des acteurs du domaine du Music Information Retrieval (MIR)

Plan

1 Reconnaissance des émotions

2 Données

3 Méthodes et traitements

Paradigmes de représentation des émotions

A Circumplex Model of Affect

James A. Russell

University of British Columbia, Vancouver, Canada

Factor-analytic evidence has led most psychologists to describe affect as a set of dimensions, such as displeasure, distress, depression, excitement, and so on, with each dimension varying independently of the others. However, there is other evidence that rather than being independent, these affective dimensions are inter-related in a highly systematic fashion. The evidence suggests that these inter-relationships can be represented by a spatial model in which affective concepts fall in a circle in the following order: pleasure (0°), excitement (45°), arousal (90°), distress (135°), displeasure (180°), depression (225°), sleepiness (270°), and relaxation (315°). This model was offered both as a way psychologists can represent the structure of affective experience, as assessed through self-report, and as a representation of the cognitive structure that laymen utilize in conceptualizing affect. Supportive evidence was obtained by scaling 28 emotion-denoting adjectives in four different ways: Ross' technique for a circular ordering of variables, a multidimensional scaling procedure based on perceived similarity among the terms, a unidimensional scaling on hypothesized pleasure-displeasure and degree-of-arousal dimensions, and a principal-components analysis of 343 subjects' self-reports of their current affective states.

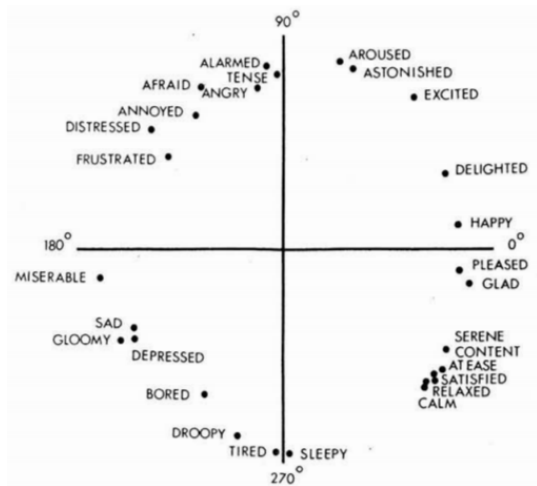
1

¹J. A. Russell. "A circumplex model of affect". In: *Journal of Personality and Social Psychology* 39 (1980), pp. 1161–1178.

Modèle de Russell (psychologie)

- En psychologie, le terme valence est utilisé pour désigner la qualité intrinsèquement agréable ou désagréable d'un stimulus ou d'une situation (axe horizontal).
- le terme activation représente le niveau d'excitation corporelle qui correspond dans notre cas au rythme musical (axe vertical)

Modèle de Russell (psychologie)



2

²J. A. Russell. "A circumplex model of affect". In: *Journal of Personality and Social Psychology* 39 (1980), pp. 1161–1178.

Plan

1 Reconnaissance des émotions

2 Données

3 Méthodes et traitements

The Million Song Dataset



3

³Thierry Bertin-Mahieux et al. "The Million Song Dataset". In: *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR 2011)*. 2011.

Million Song Dataset Benchmarks

- En utilisant un content provider, 7digital en l'occurrence, les chercheurs de l'université de vienne ont pu avoir accès à des extraits des chansons répertoriées dans MSD⁴
- Calcule de nouvelles caractéristiques à l'aide des outils jAudio et Marsyas
- Nous n'avons toujours pas les tags relatifs aux émotions

⁴Alexander Schindler, Rudolf Mayer, and Andreas Rauber. "Facilitating Comprehensive Benchmarking Experiments on the Million Song Dataset." In: *ISMIR*. 2012, pp. 469–474.



- Réseau social de plus de 30 millions d'utilisateurs
- Écouter, annoter des musiques
- Donne accès aux annotations des utilisateurs entre autres, via une api
- 505,216 tracks avec au moins un tag
- 522,366 tags uniques

Plan

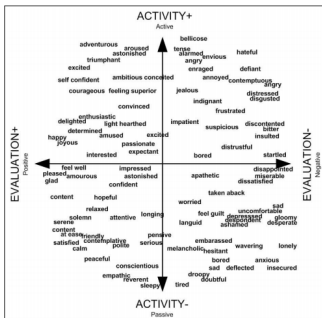
1 Reconnaissance des émotions

2 Données

3 Méthodes et traitements

Représentation spatiale des tags

- À partir des annotations des utilisateurs d'un réseau sociale musical tel que lastfm, inférer une représentation spatiale des tags relatifs aux émotions⁵⁶



⁵Cyril François Laurier et al. "Automatic Classification of musical mood by content-based analysis". PhD thesis. Universitat Pompeu Fabra, 2011.

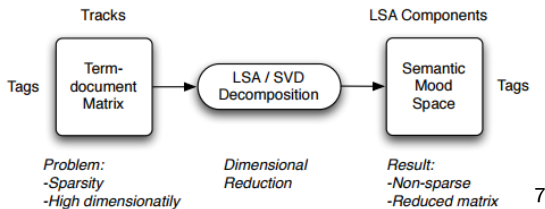
⁶Pasi Saari and Tuomas Eerola. "Semantic computing of moods based on tags in social media of music". In: *IEEE Transactions on Knowledge and Data Engineering* 26.10 (2014), pp. 2548–2560.

Term-Document Matrix

- Les tags en colonnes et les chansons en lignes
- Pour chaque chanson, les tf-idf des tags qui lui correspondent
- Résulte en une matrice creuse

$$\begin{array}{c} \text{tags} \end{array} \begin{array}{c} \text{chansons} \\ \left[\begin{array}{ccccc} x_{11} & x_{12} & x_{13} & \dots & x_{1n} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{d1} & x_{d2} & x_{d3} & \dots & x_{dn} \end{array} \right] \end{array}$$

LSA & SVD



- Choix d'un rang k pour l'approximation
- La réduction de dimensionalité doit permettre de fusionner les dimensions associées à des tags qui ont la même signification.

⁷Cyril François Laurier et al. "Automatic Classification of musical mood by content-based analysis". PhD thesis. Universitat Pompeu Fabra, 2011.

Distances entre termes

- $d_{cos}(scary, fun) = 0.99$
- $d_{cos}(tense, serene) = 0.98$
- $d_{cos}(anger, aggressive) = 0.06$
- $d_{cos}(calm, relaxed) = 0.03$

Et maintenant!



Croisement avec les données de tuvien

Table 2: Overview on features extracted from the MSD samples. *Dim.* denotes the dimensionality, *Deriv.* derivatives computed from the base features

#	Feature Set	Extractor	Dim	Deriv.
1	MFCCs [12]	MARSAYS	52	
2	Chroma [6]	MARSAYS	48	
3	timbral [13]	MARSAYS	124	
4	MFCCs [12]	jAudio	26	156
5	Low-level spectral features [11] (Spectral Centroid, Spectral Roll-off Point, Spectral Flux, Compactness, and Spectral Variability, Root Mean Square, Zero Crossings, and Fraction of Low Energy Windows)	jAudio	16	96
6	Method of Moments [11]	jAudio	10	60
7	Area Method of Moments [11]	jAudio	20	120
8	Linear Predictive Coding [11]	jAudio	20	120
9	Rhythm Patterns [9]	rp_extract	1440	
10	Statistical Spectrum Descriptors [9]	rp_extract	168	
11	Rhythm Histograms [9]	rp_extract	60	
12	Modulation Frequency Variance Descriptor [10]	rp_extract	420	
13	Temporal Statistical Spectrum Descriptors [10]	rp_extract	1176	
14	Temporal Rhythm Histograms [10]	rp_extract	420	

8

⁸Alexander Schindler, Rudolf Mayer, and Andreas Rauber. “Facilitating Comprehensive Benchmarking Experiments on the Million Song Dataset.” In: *ISMIR*. 2012, pp. 469–474.

Problème de régression

au final, la prédiction se résume en la détermination, pour une chanson donnée et un set de caractéristiques de bas et de haut niveau calculées sur celle-ci, d'un point dans l'espace des émotions obtenu auparavant.

Questions