Striani, Manuel; Vishwanath, Gautam; Díaz-Kommonen, Lily; Lieto, Antonio; Svinhufvud, Leena; Damiano, Rossana

Extracting Emotions from Users’ Annotations in Virtual Museums: a Case Study on the Pop-up VR Museum of the Design Museum Helsinki

Published in:
CEUR WORKSHOP PROCEEDINGS

Published: 06/11/2023

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.
Extracting Emotions from Users’ Annotations in Virtual Museums: a Case Study on the Pop-up VR Museum of the Design Museum Helsinki

Manuel Striani\textsuperscript{1,2,\ast,†}, Gautam Vishwanath\textsuperscript{3,‡}, Lily Diaz-Kommonen\textsuperscript{3,‡}, Antonio Lieto\textsuperscript{4,5,‡}, Leena Svinhufvud\textsuperscript{6,‡} and Rossana Damiano\textsuperscript{1,‡}

\textsuperscript{1}Computer Science Department, University of Turin, Turin 10149, Italy
\textsuperscript{2}Computer Science Institute DiSIT - University of Piemonte Orientale, Alessandria 15121, Italy
\textsuperscript{3}Department of Art and Media, Aalto University, Finland
\textsuperscript{4}University of Salerno, DISPC, Italy
\textsuperscript{5}ICAR-CNR, Palermo, Italy
\textsuperscript{6}Design Museum of Helsinki, Finland

Abstract

The paper presents a combined approach to knowledge-based emotion attribution and classification of cultural items employed in the H2020 EU project SPICE (Social cohesion, Participation, and Inclusion through Cultural Engagement)\textsuperscript{1}. In particular, we describe an experimentation conducted on a selection of items contributed by the virtual museum (Pop-up VR Museum) of Finnish design objects, created by the Design Museum Helsinki in cooperation with the Aalto University. The results show an overlapping between the emotional labels extracted from the user-generated stories attached to the objects in the collection and the emotional annotations created by the audience during the virtual visit of the collection.

Keywords

Affective Computing, Description Logics, Explainable AI, Citizen Curation, Commonsense Reasoning, VR-Museum

1. Introduction

For centuries, the role of emotions in the experience of art has been acknowledged by aesthetics; only recently, however, the availability of tools for measuring human emotions at the
physiological level has confirmed this intuition, showing that correlates of emotions, such as brain response and face expressions, are affected by the experience of art [1, 2]. In addition to their role in the way people relate to artworks [3], emotions provide a universal language through which people communicate their experience, well beyond words. An example of the capability of emotions to provide a universal means of expression is given by the diffusion of emojis, widely used also by the communities of users who may have difficulties in producing written text, such as the d/Deaf [4]. Rooted in evolution, emotion are characterized by an universal basis [5], despite the differences in their expression across languages and the cultures. In this sense, emotions can provide a way for connecting people who belong to different groups, intended as culture, age, education, and different sensory characteristics. The expression of emotions through language, in particular, lies at the basis of several models of emotions, including Shaver’s [6] and Plutchik’s [7], and has prompted the creation of a number of resources for sentiment analysis [8, 9, 10]. The application of these resources to art is straightforward: for example, WikiArt Emotions [11] is a dataset of 4,105 artworks from WikiArt annotated for the emotions evoked in the observer. The artworks were annotated via crowdsourcing for one or more of twenty emotion categories, in English language. Experiments such as WikiArt Emotions have paved the way to the extraction of emotions from text and tags to create affective art recommenders, like ArsEmotica [12] or DEGARI [13, 14, 15], able to classify and group artistic items well beyond the standard 6 basic emotions of Ekman’s theory [5], embracing richer, finer-grained models. A recent experiment on emotions evoked by art was performed in the Art Emotions Map project: 1,300 people were asked to describe how 1,500 paintings make them feel by choosing from different words. The results revealed 25 different emotions that people linked to the artworks they saw. The authors plotted these feelings on an interactive map, grouping artworks that triggered specific emotions.

In the context of the EU project SPICE (Social Participation and Inclusion through Cultural Engagement [16], which aims at supporting citizens in creating and sharing their own interpretation of artworks by attaching personal responses and affective annotations to artworks, our work has been focused on developing knowledge-based and reasoning technologies that leverage the role of emotions in the tasks of interpreting and reflecting on museums exhibits. In particular, we have developed a strategy to equip museum exhibits with emotional labels from user-generated comments that rests on the use of reasoning tools on a well-established emotion models (the Plutchik’s theory, described below). In this paper, the emotional labels have been derived from workshops and events wherein the participants listened to audio-recorded stories while testing the Pop-up VR Museum and in response, annotated their emotional feelings on a selection of items in the virtual museum (Pop-up VR Museum) in Helsinki. By applying our strategy on the user-generated contents in the Pop-up VR Museum, we obtained a fine-grained, comprehensive account of the emotions evoked by the items in the collection.

2. The Plutchik’s Ontological Model
The reference theory for the two systems is encoded in an ontology of emotional categories based on Plutchik’s psychological model of emotions [17]. The ontology structures emotional categories in a taxonomy, which currently includes 32 emotional concepts. The design of the taxonomic structure of emotional categories, of the disjunction axioms and of the object and
data properties mirrors the main features of Plutchik’s circumplex model. As mentioned before, such model can be represented as a wheel of emotions (see Figure 1) and encodes the following elements:

- **Basic or primary emotions:** Joy, Trust, Fear, Surprise, Sadness, Disgust, Anger, Anticipation; in the color wheel, this is represented by differently colored sectors.
- **Opposites:** basic emotions can be conceptualized in terms of polar opposites: Joy vs Sadness, Anger vs Fear, Trust vs Disgust, Surprise vs Anticipation.
- **Intensity:** each emotion can exist in varying degrees of intensity; in the wheel, this is represented by the vertical dimension.
- **Similarity:** emotions vary in their degree of similarity to one another; in the wheel, this is represented by the radial dimension.
- **Complex emotions:** a complex emotion is a composition of two basic emotions; the pair of basic emotions involved in the composition is called a *dyad*. Looking at the Plutchik wheel, the eight emotions in the blank spaces are compositions of similar basic emotions, called *primary dyads*. Pairs of less similar emotions are called *secondary dyads* (if the radial distance between them is 2) or *tertiary dyads* (if the distance is 3), while opposites cannot be combined.

Within this ontology, the class Emotion is the root for all the emotional concepts. The Emotions hierarchy includes all the 32 emotional categories as distinct labels. In particular, the
Emotion class has two disjoint subclasses: BasicEmotion and ComplexEmotion. Basic emotions of the Plutchik model are direct sub-classes of BasicEmotion. Each of them is specialized again into two subclasses representing the same emotion with weaker or stronger intensity (e.g. the basic emotion Joy has Ecstasy and Serenity as sub-classes). Therefore, we have 24 emotional concepts subsumed by the BasicEmotion concept. Instead, the class CompositeEmotion has 24 subclasses, corresponding to the primary (Love, Submission, Awe, Disapproval, Remorse, Contempt, Aggressiveness and Optimism), secondary (Hope, Guilt, Curiosity, Despair, Unbelief, Envy, Cynicism and Pride) and tertiary (Anxiety, Delight, Sentimentality, Shame, Outrage, Pessimism, Morbidness, Dominance) dyads. Other relations in the Plutchik model have been expressed in the ontology by means of object properties: the hasOpposite property encodes the notion of polar opposites; the hasSibling property encodes the notion of similarity and the isComposedOf property encodes the notion of composition of basic emotions.

3. DEGARI

The core component of DEGARI relies on a probabilistic extension of a typicality-based Description Logic called $\mathcal{T}^cL$, (Typicality-based Compositional Logic, introduced in [18]). This framework allows one to describe and reason upon an ontology with commonsense (i.e. prototypical) descriptions of emotional concepts, as well as to dynamically generate novel prototypical concepts in a knowledge base as the result of a human-like recombination of the existing ones [19, 20, 21, 22, 23, 24].

The logic $\mathcal{T}^cL$, that we recall here for self-containedness, is the result of the integration of two main features: (i) an extension of a nonmonotonic Description Logic of typicality $\mathcal{ALC} + \mathcal{T}_R$ introduced in [25] with a distributed semantics; (ii) a well-established heuristics inspired by cognitive semantics for concept combination and generation [26], in order to formalize a dominance effect between the concepts to be combined: for every combination, it distinguishes a HEAD, representing the stronger element of the combination, and a MODIFIER. The basic idea is to extend an initial knowledge base (ontology) with a prototypical description of a novel concept, obtained by the combination of two existing ones, namely a HEAD concept and a MODIFIER concept. In this logic, typical properties can be directly specified by means of a typicality operator $\mathcal{T}$ enriching the underlying Description Logic, and a knowledge base can contain inclusions of the form $p : : \mathcal{T}(C) \sqsubseteq D$ to represent that "typical Cs are also Ds", where $p$ is a real number between 0.5 and 1, representing the probability of finding elements of $C$ being also $D$. From a semantic point of view, it considers models equipped by a preference relation among domain elements, where $x < y$ means that $x$ is “more normal” than $y$, and that the typical members of a concept $C$ are the minimal elements of $C$ with respect to this relation. An element $x$ is a typical instance of a given concept $C$ if $x$ belongs to the extension of the concept $C$, written $x \in C^{\mathcal{T}}$, and there is no element in $C^{\mathcal{T}}$ "more normal" than $x$. $\mathcal{T}^cL$ also considers the key notion of scenario. Intuitively, a scenario is a knowledge base obtained by considering all rigid properties as well as all ABox facts, but only a subset of typicality properties. To this aim, it considers an extension of the Description Logic $\mathcal{ALC} + \mathcal{T}_R$ based on the distribution semantics known as DISPONTE [27]. The idea is to assume that each typicality inclusion is independent from each other in order to define a probability distribution over scenarios: roughly speaking,
a scenario is obtained by choosing, for each typicality inclusion, whether it is considered as true of false. Reasoning can then be restricted to either all or some scenarios. $T^{CL}$ equips each scenario with a probability, easily obtained as the product, for each typicality inclusion, of the probability $p$ in case the inclusion is involved, $(1 - p)$ otherwise. It immediately follows that the probability of a scenario introduces a probability distribution over scenarios, that is to say the sum of the probabilities of all scenarios is 1.

The bridge from the definition of the emotions in the ontology and the annotations associated with an artwork is provided by an emotion lexicon. Emotional concepts are described by using the NRC Emotion Intensity Lexicon [8] (one of the lexica used also by SOPHIA). Such lexicon provides a list of English words, each with real-values representing intensity scores for the eight basic emotions of Plutchik’s theory. The lexicon contains close to 10,000 words, including terms already known to be associated with emotions as well as terms that co-occur in Twitter posts that convey emotions. The intensity scores were obtained via crowdsourcing, using best-worst scaling annotation scheme. In this work, we considered the most frequent terms available in such lexicon (and associated to the basic emotions of the Plutchik wheel) as typical features of such emotions. In this way, once the prototypes of the basic emotional concepts were formed, the $T^{CL}$ reasoning framework was used to generate the compound emotions.

In the context of our system, $T^{CL}$ allows us to provide a formal, explainable framework for combining prototypical descriptions of concepts. It is adopted to automatically build the prototypical representations of the compound emotions according to the Plutchik’s theory. The prototypes of basic emotions are formalized by means of a $T^{CL}$ knowledge base, whose TBox contains both rigid inclusions of the form

$$\text{BasicEmotion} \sqsubseteq \text{Concept},$$

and constraints, e.g. $\text{Joy} \sqsubseteq \text{PositiveEmotion}$ as well as prototypical properties:

$$p : : T(\text{BasicEmotion}) \sqsubseteq \text{TypicalConcept},$$

representing typical concepts of a given emotion, where $p \in (0.5, 1]$, expressing the frequency of such a concept in items belonging to that emotion: for instance, $0.72 : : T(\text{Surprise}) \sqsubseteq \text{Delight}$ is used to express that the typical feature of being surprised contains/refers to the emotional concept Delight with a frequency/probability/degree of belief of the 72%.

Once the association of lexical features to the emotional concepts in the Plutchik’s ontology is obtained and the compound emotions are generated via the logic $T^{CL}$, the system is able to reclassify the artworks in the novel emotional categories. Intuitively, an item belongs to the new generated emotion if its metadata (name, description, title, user-generated annotations) contain all the rigid properties as well as at least the 30% of the typical properties of such a derived emotion. The 30% threshold was empirically determined: i.e., it is the percentage that provides the better trade-off between overcategorization and missed categorizations [28].

4. Analyzing the Stories from Pop-up VR-Museum

The Pop-up VR Museum is a collaborative endeavor between Aalto University and Design Museum Helsinki (DMH), resulting in a prototype of a Virtual Reality (VR) application that
provides citizens with access to a digitized collection of artifacts from DMH. Some of the features offered by the application include:

- **Interactive Object Manipulation**: Users have the capability to interact with virtual design objects, enabling actions such as picking them up, rotating them, and even altering their color variants within the VR environment.

- **Engagement with narratives**: Users are encouraged to contribute their own stories, while also having the opportunity to read and listen to narratives shared by other contributors. Additionally, users can annotate these stories using a range of emotions, represented through emojis.

- **Immersive Environments**: Each virtual object is accompanied by a distinct virtual environment, allowing users to immerse themselves and gain fresh perspectives on the object. This VR experience employs a Natural User Interface (NUI) to facilitate user interaction, characterized by its minimal interaction requirements (Reference needed here). This design choice results in a gentle learning curve for users, potentially enhancing their sense of presence within the VR environment as well (References needed).

In the Pop-up VR Museum, the functionality for annotating emotions becomes available to users exclusively after they have thoroughly engaged with the content. Specifically, this feature is activated once a user has completed the process of listening to a narrative and immersing themselves in the design object. This criterion was deliberately chosen in the case study to guarantee that users experience the entirety of a story without any premature interruptions. Moreover, it ensures that users have the opportunity to engage with the object from a novel perspective within the VR environment. It is only under these circumstances that a user is deemed eligible to select and annotate their emotions to a particular story.

Before the integration of DEGARI 2.0 emotions for user annotation, the initial iterations of the prototype employed a Likert scale range (1-5) of commonly used emojis, as typically seen in surveys for gathering user feedback. This range encompassed the following emotions, ranging from sadness to happiness: 😞, 😢, 😡, 😃, and 😊. Notably, these emojis were not explicitly labeled for users, leaving them open to individual interpretation. Users were limited to selecting a single emoji, and the experience would proceed accordingly.

In the subsequent iterations, a more nuanced approach was introduced by offering users the ability to annotate their emotions using the DEGARI 2.0 framework, which comprises nine distinct and complex emotional states. Users were presented with a specific question: "How did the previous story make you feel? (Select multiple)", accompanied by the option to choose multiple emotions that resonated with them. In order to select emojis that closely aligned with each DEGARI 2.0 emotion, the following were chosen:

- Delight: 😄
- Love: 😍
- Joy: 😃
- Optimism: ☀️
- Hope: 😊
- Curiosity: 🎨
- Disapproval: 😞
- Anxiety: 😞
- Optimism: 😊
- Outrage: 😡

However, it remains a subject of ongoing inquiry whether these emojis accurately encapsulate the full spectrum of DEGARI 2.0 emotions, especially in light of the potential language variations
introduced by users selecting either Finnish or Swedish for their Pop-up VR Museum experience. In such cases, the selected texts would subsequently be translated into the chosen language, further emphasizing the need for continued evaluation and refinement.

In the broader context, the selection of emotions is influenced by a diverse array of factors:

1. User’s Emotional Response to the Story: The emotions chosen are contingent upon how the user personally interprets and reacts to the narrative presented.
2. Content Characteristics in the Virtual Environment: The emotional selection may also be influenced by the nature of the content within the virtual environment that the user is immersed in.
3. Sense of Presence in the Experience: The degree to which users feel present in the virtual environment is a critical determinant. Any disruptions in this sense of presence may lead to users expressing frustration and subsequently annotating negative emotions to stories.
4. Physical Surroundings during the VR Experience: Users engaging with the Pop-up VR Museum while immersed in VR may also be impacted by their immediate physical environment. This is particularly pertinent as many users often engage in conversations with mediators located in the real space, potentially influencing their emotional annotations.

The fist experiment conducted with DEGARI 2.0 was on the stories of the virtual museum (Pop-up VR Museum) in Helsinki (the used dataset is on the Linked Data Hub (LDH) and has been used also for further analysis by the Thematic and Value Reasoners). In this case the term story refers to the users generated contents on specific artifacts. In particular, the first part of Table 1(a), shows some statistics calculated from the use of our DÈGARI 2.0 system for the extraction of emotions on user-generated comments. The set of tuple of emotions that have been extracted by DEGARI 2.0 most frequently are {Delight, Joy}, {Love, Optimism} and {Love, Joy}. Table 1(b) shows for each extracted emotion from user-generated comments on stories, the relative frequency. The emotion that DEGARI 2.0 extracted most frequently, was Joy with 36 followed by Delight with 25.

The second experiment was to validate DEGARI 2.0 as a cognitive tool to investigate the true emotional content that is present within the stories. In particular, we wanted to understand how the emotions extracted from DEGARI 2.0 could be aligned with the emotions chosen by the users during the creation of their stories. In particular, the fist column shows the emotions extracted by DÈGARI 2.0: the emotion "Love" was extracted 19 times out of a total of 20 objects (95%), 15 times DEGARI 2.0 extracted the emotion "Optimism" (75%), 9 times the emotion "Hope" (45%), 4 times "Joy" and "Delight" (20%). The 4th column shows the number of users who chose a particular emotion related to the 20 objects. Specifically, 17 users chose "Love", 14 "Optimism", 4 the emotion "Hope", 8 the emotion "Joy" and 7 users the emotion "Delight". Finally, the last column shows the perfect match between the emotions extracted by DEGARI 2.0 and the emotions selected by the users. In particular, 17 users have chosen the emotion "Love" (85% of them had a perfect match with the 19 extracted by DEGARI 2.0). These results are shown in Table 2 and Table 3.

1https://spice.kmi.open.ac.uk/dataset/details/104
### Table 1
Statistics for DEGARI 2.0 extracted emotions from Pop-up VR Museum Stories (a) and Frequency for DEGARI 2.0 extracted emotions (b)

(a) | Emotion     | Frequency |
---|-------------|-----------|
| Joy          | 36         |
| Delight      | 25         |
| Love         | 8          |
| Optimism     | 5          |
| Curiosity    | 1          |
| Pessimism    | 3          |
| Disapproval  | 3          |
| Hope         | 2          |
| Anxiety      | 1          |
| Shame        | 1          |
| Outrage      | 1          |

(b) | Emotion     | Frequency |
---|-------------|-----------|
| Joy          | 36         |
| Delight      | 25         |
| Love         | 8          |
| Optimism     | 5          |
| Curiosity    | 1          |
| Pessimism    | 3          |
| Disapproval  | 3          |
| Hope         | 2          |
| Anxiety      | 1          |
| Shame        | 1          |
| Outrage      | 1          |

Table 2
Aggregate statistics on the 72 created stories from 72 DMH-users on a total of 20 objects. The total number of emotions extracted by DEGARI 2.0 on all 72 stories is 51 (70.83 %)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DMH-Stories</td>
<td>72</td>
</tr>
<tr>
<td>DEGARI extracted emotions</td>
<td>52  (70.83% on total stories)</td>
</tr>
<tr>
<td>Total DMH-users</td>
<td>72</td>
</tr>
<tr>
<td>Total Objects</td>
<td>20</td>
</tr>
</tbody>
</table>

5. Discussion and Conclusion

In this paper, we described the results of the analysis on the stories generated by the users for the design artefacts exhibited in the Pop-up VR Museum created by the Aalto University and Design Museum Helsinki (DMH) using the affective-based sensemaking system called DEGARI 2.0. The results of the analysis show that there is an overlapping between the emotional labels extracted from the user-contributed stories and the emotional labels added by other users to the same objects as part of their experience with the Pop-up VR-Museum. Although the sentiment of the user contributions (stories and annotations) is generally oriented towards positive emotions – due to the very same natures of the experience, which triggers a reflection on the personal engagement with the artefacts –, we think that data extracted by DEGARI 2.0 complement the
Table 3
Statistics about DEGARI 2.0 extracted emotions, users selected emotions and intersection between them (perfect match)

<table>
<thead>
<tr>
<th>Emotions</th>
<th>DEGARI 2.0 Extracted emotions</th>
<th>用户的交集</th>
<th>Emotions selected by users</th>
<th>% of perfect match</th>
</tr>
</thead>
<tbody>
<tr>
<td>Love</td>
<td>19</td>
<td>95.00%</td>
<td>17</td>
<td>70.00%</td>
</tr>
<tr>
<td>Optimism</td>
<td>15</td>
<td>75.00%</td>
<td>14</td>
<td>70.00%</td>
</tr>
<tr>
<td>Joy</td>
<td>9</td>
<td>45.00%</td>
<td>4</td>
<td>20.00%</td>
</tr>
<tr>
<td>Delight</td>
<td>4</td>
<td>20.00%</td>
<td>8</td>
<td>40.00%</td>
</tr>
<tr>
<td>Hope</td>
<td>4</td>
<td>20.00%</td>
<td>7</td>
<td>35.00%</td>
</tr>
</tbody>
</table>

users’ annotation in a way that can provide the museum curators with an useful insight on the emotional response of the audience expressed. In particular, the overlapping between the two types of data suggests that a basic human mechanism for the creation of social cohesion, namely emotional empathy, is at work also in the experience of art, paving the way to an ethical use of emotions to build cohesion.

References


