Emotional and Engaging Movie Annotation with Gamification

Lino Nunes, Cláudia Ribeiro[a] and Teresa Chambel[b]
LASIGE, Faculdade de Ciências, Universidade de Lisboa, Portugal

Keywords: Movies, Emotions, Annotations, Gamification, Games with a Purpose, Emotional Journaling, Challenge, Explore, Annotate, Review, Compare, Achieve.

Abstract: Entertainment has always been present in human activities, satisfying needs and playing a role in the lives of individuals and communities. In particular, movies have a strong emotional impact on us, with their rich multimedia content and their stories; while games tend to defy and engage us facing challenges and hopefully achieving rewarding experiences and results. In this paper, we present a web application being designed and developed to access movies based on emotional impact, with the focus on the emotional annotation of movies, using different emotional representations, and gamification elements to further engage users in this task, beyond their intrinsic motivation. These annotations can help enriching emotional classification of movies and their impact on users, with machine learning approaches, later helping to find movies based on this impact; and they can also be collected as personal notes like on a journal where users collect the movies they treasure the most, and that they can review and even compare along their journey. We also present a preliminary user evaluation, allowing to assess and learn about its perceived utility, usability and user experience, and to identify most promising features and directions for further improvements and developments.

1 INTRODUCTION

Watching a movie, until a few years ago, was an emotional though passive experience, interacting with our emotions, awakening nostalgia, feelings and memories, but with the time and rhythm defined by the movie. This has changed with technology allowing to search for, and control what and when to watch and rewatch movies, and at what pace; but the emotional impact has hardly been addressed and supported (Horner, 2018; Oliveira et al., 2013). In order to add this dimension, movies need to be annotated or classified, based on their emotional content or the emotional impact they have on viewers. Each film can be analyzed in a wide variety of content, such as soundtracks, subtitles or images; in separate or in multimodal approaches; on the other hand, viewers emotional impact can be assessed automatically with the aid of sensors, or by self-report; and these can be done in realtime or post-stimulus. Users can annotate movies: 1) to help creating catalogues or datasets to train machine learning algorithms for a more robust automatic movie classification, in a human computation perspective (Garrity & Schumer, 2019|); or 2) they may also annotate them as a way to personalize, keep and review the user's emotional perspective of the movies they watch along time, like in a journal. They could even review and compare how they felt when they watched the same film at different times, possibly years apart, at different phases in their lives. In the latter, for personal annotations, the motivation to annotate would tend to be more intrinsic; while in the former, users might need additional extrinsic motivation, e.g. through gamification elements, like challenges, awards and achievements, adding to the pride of contributing to the movie community, especially for movies they really appreciate, care for, and sometimes know by heart.

In this paper, we present a web application being designed and developed to access movies based on emotional impact, with the focus on the emotional annotation of movies, using different emotional representations, and gamification elements to further engage users in this task, beyond their intrinsic motivation. We also present a preliminary user evaluation, allowing to assess and learn about its

[a] https://orcid.org/0000-0002-1796-0636
[b] https://orcid.org/0000-0002-0306-3352

Nunes, L., Ribeiro, C. and Chambel, T.
Emotional and Engaging Movie Annotation with Gamification.
DOI: 10.5220/0010991500003124
In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022) - Volume 2: HUCAPP*, pages 262-272
ISBN: 978-989-758-555-5; ISSN: 2184-4321

perceived utility, usability and user experience, and to identify most promising features and directions for further improvements and developments.

2 BACKGROUND

This section presents main concepts and related work, to contextualize our own work and contributions.

2.1 Emotions

There is no single definition for emotions but we can retain this one, used in psychology: emotions are "defined as a state of believe which results in psychological changes" (Sreeja & Mahalakshmi, 2017) and human beings are constantly demonstrating their emotions, but representing and annotating becomes complicated, as there are too many and their definition is not always clear and consensual. It is from this difficulty that the need to create models for their representation arises.

Categorical models describe emotions as discrete categories like Ekman's (1992) basic emotions (Anger, Fear, Sadness, Happiness, Disgust and Surprise), or through a scale that classifies them as positive, negative or neutral (Choi & Aizaha, 2019). Dimensional models represent emotions in space, in 2-3 dimensions, like Russell's (1980) and Plutchik's (1980); where categorical emotions can be located.

Russell's circumplex model represents emotions in a two-dimensional space (VA) based on: Valence, as a range of positive and negative emotions; and Arousal, representing their level of excitement (Sreeja & Mahalakshmi, 2017). Plutchik (1980) defined a 3D model, both categorical and dimensional (polarity, similarity, intensity), with 8 primary emotions: 6 like in Ekman's, plus: anticipation, and trust, represented around the center, in colors, with the intensity as the vertical dimension (in 3 levels).

2.2 Affective Computing & Journals

Providing support for emotions through systems and devices that can recognize, interpret, process, and simulate human affect is the main goal of Affective Computing (Picard, 2000). The related field of Positive Computing [3], on the other hand, informs the design and development of technology to support psychological wellbeing and human potential; in 3 approaches: 1) preventative, when technology is redesigned to address or prevent detriments; 2) active, to consider and promote the wellbeing of individuals; and 3) dedicated, where technology is created and totally dedicated to promoting wellbeing. We adopted the 2nd (active), by incorporating emotional support in a system aimed at watching movies of all kinds.

The personal journal concept, in the digital age, could be an example of the 3rd approach. A personal journal can be e.g. a notepad, a diary, a planner, a book; and in digital formats they can include media, allowing users to organize their goals and "to do"s, collect quotes, and register their own notes, describing what happened, interested them or made them happy (Garrity & Schumer, 2019; Chambel & Carvalho, 2020). According to Ryder Carrol, the creator of the bullet journal method, while flexible and personalized, it helps focusing on what's important. In the positive computing perspective, personal journals encourage and favor an attitude of self-awareness and mindfulness, that may contribute to users' wellbeing.

2.3 Emotional Impact Annotation

Annotation can provide the means for media classification and personal notes or journals; and they can be made in realtime (often with continuous methods) allowing to capture the temporal nature of the emotions, or post-stimulus (often with a discrete scale).

The most well-known discrete emotional self-report tool is probably the Self-Assessment Manikin (SAM). It consists of manikins expressing emotions and varies along three dimensions: arousal, valence and dominance or control (Bradley and Lang, 1994; see section 3.2.2). Valence varies from negative to positive, Arousal varies from calm to excited, and Dominance from low to high sense of being in control. Its predecessor, SDS: Semantic Differential Scale, was based on words and less flexible or accessible to non-native english speakers.

As continuous annotation approaches: FeelTrace (Cowie et al., 2000) uses mouse or joystick for continuous input and for output, based on Russell's circumplex. Plutchiks colors are used for the extremes in the axes, and emotions are represented by colored circles with the color interpolated by those on the axes, that get smaller as new ones get in. DARMA is based on CARMA (that allows annotations in each dimension in separate), allowing annotations in 2D, choosing both dimensions in just one selection, using a joystick (Girard & Wright, 2018). It also presents a line chart of evolution, and compares with previous annotation, estimating agreement and confidence.

Zhang et al. (2020) describe a tool for emotional annotations in videos on mobile devices, choosing VA with a virtual joystick represented by a circle as foreground to a wider circle for the circumplex (at the

lower right corner of the screen), with a color for each quadrant; colors highlighted at the border of the video, when annotations occur; and more transparent in the center.

The works above were focused on emotional annotation support, possibly for movie watching, but not so specific. In a perspective of expressing emotions, closer to the concept of a personal journal, although not for movie, we highlight: In MoodMeter (http://moodmeterapp.com/process/), users choose a color and see which emotions are related, make short descriptive reflections on what they are feeling, and are offered regulatory strategies using quotes, images and practical tips; and can revisit felt emotions at any time Not focused on media. PaintMyEmotions is an interactive self-reflection system where users express their emotions through painting (facilitates expression), photography (helps arousing emotions) and writing (helps expression and clarification) (Nave et al., 2016). And in Cove (Humane, 2015-19), users create small loops of music to help them express how they feel in a safe, positive environment, in the perspective of the creator, not as content annotation.

2.4 Gamification & Engagement

Games can be merely playful and entertaining, or they may be designed as engaging serious games, for other purposes, like health or education (Plaisent et al., 2019). And these have existed even before the digital era in most societies, e.g. teaching in ways that were also playful. Computer games design is based on psychology, to account for motivations and emotions; and the more technical areas of gamification and programming (Huotari & Hamari, 2012).

Human needs were represented by psychologist Maslow in a hierarchy (Maslow, 1943; McLeod, 2020), with a priority ranging from: the most basic, related with survival; up to the most complex, like self-fulfillment, which can often be satisfied with the help of intrinsic motivations, when we are moved by the will and interest in some activity, and the reward itself is this behavior (Pink, 2011).

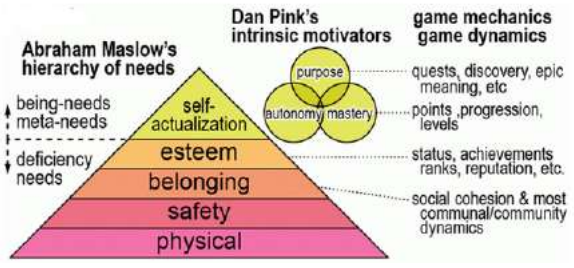

Figure 1: Maslow's pyramid of needs vs. Pink's intrinsic motivators vs. game dynamis (Wu, 2011).

The elements and dynamics of games tend to use these needs and intrinsic motivations to become attractive and effective means of entertainment (Fig.1). With the use of flow theory (Csikszentmihalyi, 2014), it is possible to further understand how players relate to challenges proposed in the games.

Gamification can be defined as the use of game design elements in non-game contexts (Deterding et al., 2011). Whereas, Human computation is defined by Law & Ahn (2009) as: “the idea of using human effort to perform tasks that computers cannot yet perform, generally in a pleasant way” .

Serious games like Google Image Labeler and Phetch are some of the applications that use the concepts of human computation and gamification to enrich image search systems. Google Image Labeler was aimed at categorizing images, to improve their databases and searches (Chitu, 2016). It combines two random users who receive the same set of images, to be annotated in two minutes with the maximum number of labels (and prohibited words); and for the two users to receive points, their labels must match and the amount depends on how specific the label is for the image. Ranking tables are used, teams with high scores can enter one or two ranking tables with the names Today's Top Pair and All-time Top Contributors. The Phetch online game (Ahn et al., 2007) can be played by three to five players; one the describer, the others are the seekers who receive the description from the describer and have to interpret that description to find the correct image. The describer and seekers do not communicate; seeker discovers, both win; and possible to play with a bot.

With a similar goal, but another media type, TagATune is aimed at classifying music excerpts (Law et al., 2007). It involves two users, who are given thirty-second audio clips, which can be the same or different. To make its description easier, the songs used are popular and easily recognizable by players (Law & Ahn, 2009). Using tags, each player must describe their audio snippet, then try to identify if snippets are the same. This game has a scoring system and a ranking system, and players only receive points if they both get the right music snippet, with no penalty if they pass a snippet of music they don't rate.

iHEARuPLAY was built primarily for annotating audio, but it can also incorporate images and video databases. Several gamification elements were used to intrinsically motivate users (Hantke et al., 2015), it is based on a credibility system, and it also shows the multiplier that is used to calculate the score.

Waisda? is an application to annotate video, based on the fundamental ideas of the serious game ESP (Hildebrand, et al., 2013), and as in previous ones,

uses tags, rewards agreement among participants, and presents leaderboards; and uses tagclouds for most used tags (we use: to represent emotional overviews or summary, based on categorical emotions), and it allows more than two users per session

Tag For Video (TAG4VD) is a video annotation application with the help of users. They individually annotate videos using tags and evaluate them to help describe videos, thus creating metadata (Viana & Pinto, 2017). Users can interact with the application as guests/view or in competitive mode, where they have challenges and rewards, can annotate videos, and are rewarded with some points. It uses tagclouds and users can mark parts of the video to annotate.

Most of these systems are aimed at annotating content, where a "right" answer tends to exist. That is not so true when annotating emotional impact; but still, gamification elements can add to the intrinsic motivation to belong to the community of movie fans, to be challenged and achieve; and on the other hand, given the subjectivity, it gets closer to the concept of the personal journal.

3 EMOTIONAL ANNOTATION WITH GAMIFICATION

This section presents the main features for emotional and engaging annotation of movies with gamification in the AWESOME web application, being designed and developed to access movies based on emotions.

3.1 Gallery View

When users first access the application they are prompted with a login page, where they can register or login. The first page is the Gallery View (Fig. 2), with two lists of movies: 1) movies that the user can choose to annotate; or 2) "Continue To Annotate" where they left of.

3.2 Annotation View

This view allows creating emotional annotations in movies. Is accessed from the menu Annotation (upper right corner above the video player) and is composed by: the emotion wheel, self-assessment manikin, and categorical emotion annotation interfaces, and the timeline, further described in the following sections.

3.2.1 Emotion Wheel Annotation

For the emotion wheel annotation, we draw on Russell's Circumplex model of emotion (1980), where each annotation component is designed according to the valence and arousal dimensions.

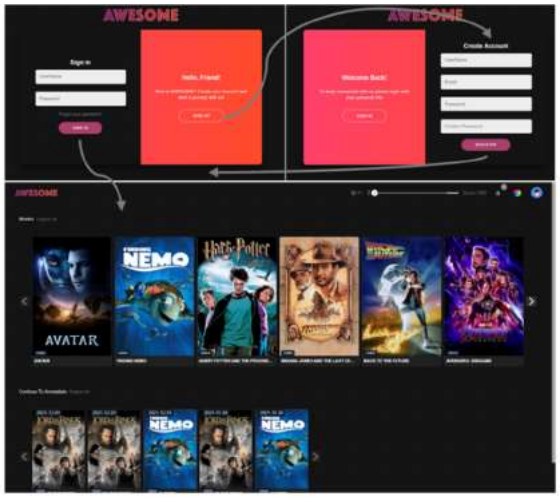

Figure 2: (a) login; (b) registration pages; (c) Gallery View.

To accomplish this, we implemented a virtual joystick (Fig. 3), with arousal in the vertical axis, and valence in the horizontal axis. The use of a joystick as an emotion input method is considered to be advantageous as it allows for continuous and simultaneous acquisition of valence-arousal (VA) annotations.

To provide a visual cue and help users to identify the emotion around the wheel, the virtual joystick is represented by a colour palette similar to Plutchik's (1980) or Geneva (Sherer, 2005) wheel of emotions. These as default options to choose (on the top right), but it will be possible to add different colour palettes.

The virtual joystick (for input) as well as the analogous feedback wheel above (for output) starting with no colour and accumulating the colors in a path along time, are placed top right in the screen (Fig. 3). The radius of both the virtual joystick as well and the feedback wheel, are automatically adjusted to the size of the screen. A gradual transparency was also applied to ensure minimal occlusion.

To use the annotation tool, users need to place the mouse pointer on the virtual joystick for inputting their valence and arousal. Two modes are supported: a unique click on the wheel leads to a unique VA value captured and associated to the current timestamp of the movie; dragging the mouse pointer without releasing it along the wheel leads to a realtime continuous VA capture. Sampling rate is 40 milliseconds, which guarantees that a VA is acquire per video frame. Realtime continuous mode allows to capture emotional states along movie scenes.

Two types of affective feedback are provided: a colour frame on the top of the video and the feedback

wheel. The colour frame shows the current colour, while the user is annotating; in between annotations, it corresponds to the last colour added by the user.

The feedback wheel shows, in realtime, the paths the user is doing while dragging the mouse pointer. For each path position, a circle is added with the respective colour and an initial size; if the user stays in the same position for a period of time, the size of the circle is adjusted accordingly.

Each annotation is added to the timeline, where single annotations are represented by a coloured stick, while realtime continuous annotations are represented by a larger colour gradient with colours' percentage proportional to the duration of those emotions.

Figure 3: Annotation View – Emotion Wheel Annotation (a) colour frame; (b) feedback wheel; (c) virtual joystick; d) example of single annotation in the timeline; and, e) example of realtime continuous annotations in the timeline.

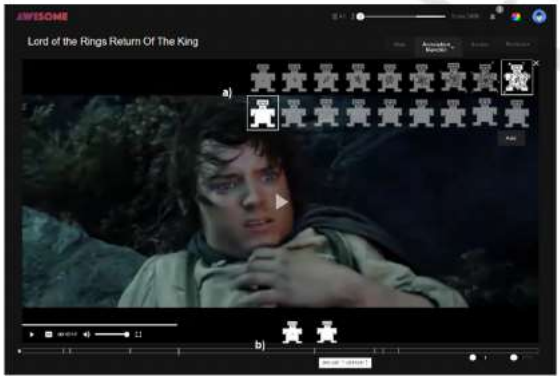

Figure 4: Annotation View – Self-assessment Manikin Annotation (a) 9-point scale represented in a Grid; (b) feedback provided when hovering over an annotation.

3.2.2 Self-assessment Manikin Annotation

We are using the first two dimensions of SAM: Valence and Arousal, for being the ones found in Russell's circumplex, the model inherent to the emotion wheel. For the creation of annotations using SAM (in section 2.3), two scales were implemented, for flexibility: 5 and 9-point. When users start an annotation session, they can choose which scale to use, and it will be maintained throughout the session. Similar to the wheel annotation, the SAM manikins are positioned on the top right side of the screen, organized in a grid with top row for valence and bottom row for arousal (Fig. 4). Initially the manikin grid is presented with a transparency and it becomes opaque as the user chooses the correspondent valence and arousal representation. Each annotation is added to the timeline, and on hover, the manikins as well as valence and arousal values are shown.

3.2.3 Categorical Emotion Annotation

This annotation is based on a set of categorical emotions, with more positive emotions than Ekman's (1992) which only has one positie in six (happy). We designated a colour to each category. This was accomplished by attributing a valence and arousal value to each category and finding out on the colour palette the corresponding colour at those coordinates. The representation and functioning of the categorical emotion annotation is very similar to the SAM annotation. It is positioned on the top right side of the screen, organized in a grid. Each annotation is added to the timeline and on hover, it shows the text label with the corresponding emotion.

3.2.4 Timeline

The timeline supports two modes: instants and segments (Fig. 5). In the instant mode, the timeline has no bounds, allowing the user to freely annotate movie instants. An instant can correspond to a movie frame or a sequence of movie frames (as preconfigured). In the segment mode, the timeline is divided by movie segments, that can be of equal duration, or previously identified scenes of interest, with different durations. In the segment mode, each annotation type can be used to annotate each scene as a post-stimulus discrete self-report, with only one annotation (of each type) per scene.

The timeline supports zooming on a particular annotation as well as in the timeline itself, allowing the user to more thoroughly inspect the annotations. A player head was integrated in the timeline, allowing the user to easily identify and accompany the movie progression. It also supports point and click to skip to a different movie time, and hovering through the annotations to inspect affective feedback (Fig. 3-4). In the case of a realtime continuous wheel annotation, it highlights the annotation path in the feedback wheel corresponding to that time. While hovering an

annotation, the user can use the mouse right button to delete the annotation. Finally, the user can choose to only have the current annotation type timeline, or the three annotation types timelines simultaneously: more information, for more control and comparison.

Figure 5: Timeline component (a) Instant timeline; (b) view timeline for all annotation type; (c) Segment timeline; and, (d) change to segment annotation.

Figure 6: Movie Detail View data visualization – from top to bottom, wheel emotion annotation, SAM emotion annotation, and categorical emotion annotation.

3.3 Movie Detail View

The Movie Detail view is accessed when the user chooses a movie from Movie Gallery (Section 3.1). It presents two types of information: 1) movie details, including the IMDB ranking, number of reviews, year, movie duration, a summary of the synopsis and the cast; and 2) an aggregate view of all the annotations produced by the users, that can be represented in different ways: (a) emotion wheel, (b) SAM scale, and (c) a tag cloud with the categorical emotions. The user can switch between these forms of representation by pressing the buttons representing each type of annotation, located next to the movie time (Fig. 6). When this page is accessed, it shows as default the emotion wheel, with all its annotations. If the user changes the annotation to type SAM, the SAM 9-point scale with the selected valence and arousal images is presented. Although the application also supports SAM 5-point scale (for flexibility), in this view, we chose the 9-point because it is the super set, including all pictorial images of the 5-point scale. The highlighted images correspond to the statistical mode (the most frequent) for V and A, aggregating all the annotations produced with the SAM. The tag cloud shows a set of emotion categories, with the size of the words/emotions reflecting their frequency in the categorical annotations made by the user. When we hover a word, it shows the number of times that emotion appears in the annotations.

Finally, for each annotation type, there is a timeline presented, to contextualize annotations along time. All the features described in sub-section 3.2.3 are supported here, synchronizing the timeline with the other representations, except for delete (in this summary view, one can view & explore, not change).

3.4 Emotional Journal and ReView

The Emotional Journal (Fig. 6), shows the emotional information for each movie previously annotated, and it can be accessed from the Profile menu. Each movie is represented in a list by the movie cover and aggregated emotional information created with: the emotion wheel, the self-assessment manikin, and the categorical emotion (sections 3.2.1-3). It also has the date the movie was first annotated and last updated.

Figure 7: Emotional Journal view: a) Profile view; b) Emotional Journal; c) ReView annotation sessions; and, d) create new annotation session.

For each annotated movie, the user can create a new annotation session, or review and compare annotation sessions. And this list can be filtered by movie title.

In this ReView (Fig. 8) the user can review or compare annotation sessions having the annotations being presented synchronously while the user rewatches the movie. Each annotation type can be presented in separate or simultaneously. When the annotations are reviewed by type, they are shown in the same representation type used to make them, like in the movie detail (Fig. 6). When annotation types are reviewed simultaneously, all the annotations are presented at once; but to reduce complexity, the user can choose a common desired output representation: emotion wheel, SAM grid, or tag cloud; since they can be converted through their underlying VA value (rounded up when the output precision is lower). To represent a categorical emotion annotation in a wheel, we use that emotion's VA values. To represent the wheel and SAM annotations in the tag cloud, we use the Euclidean distance of the VA values, to find out the closest categorical emotion for each annotation – the ones that will be represented in the tag cloud.

Users can also compare two annotation sessions. In Fig.8, comparing SAM annotations from session 5, with wheel annotations from session 1. Finally, sessions can be filtered by creation date, annotation method (annotation or sensor), type (wheel, SAM, categorical) and interval (instant or segment).

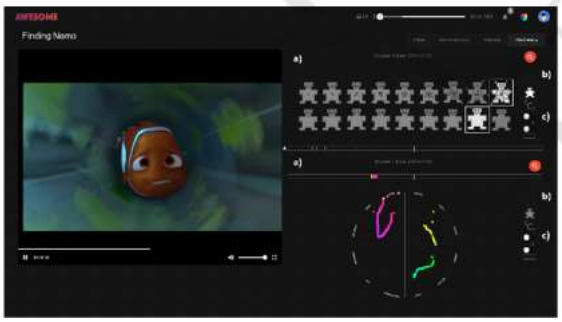

Figure 8: ReView view: a) movie sessions; b) annotations representations – emotional wheel, 9-point scale SAM, and tag cloud; c) view all annotation types simultaneously, and change to timeline segments.

3.5 Gamification

Gamification can be defined as the use of game design elements in non-game contexts (Deterding et al., 2011). Game elements can be designed to augment and complement the entertaining qualities of movies, motivating and supporting users to contribute to the content classification, combining utility and usability aspects (Deterding et al., 2011; Khaled, 2011). Main properties to aim for include: Persuasion and Motivation, to induce and facilitate masscollaboration, or crowdsourcing; Engagement, possibly leading to increased time on the task; Joy, Fun and improved user experience; Reward and Reputation inspired in incentive design. In the following sections we describe the gamification elements and how they are connected to these main properties.

3.5.1 Challenges and Achievements

In order to persuade and motivate users, we implemented two different types of challenges: time constraint and daily. The daily challenge is a task that is proposed every day, and can be an annotation task related to a movie a user has seen often, or similar movies. Time constraint challenges are tasks that the system gives the user to complete in a predetermined time frame. Whenever a challenge is completed, the user receives rewards, which can be achievements, points, or help, to motivate the user in the process of annotating movies. The achievements can be interpreted as a sticker book in which users must complete tasks to unlock them, as they progress. These two gamification elements can be accessed via the user profile (Fig. 8) and the notifications menu.

3.5.2 Points and Levels

We implemented a points system as a strategy to engage, and possibly lead the user to increase time on task, specifically annotating movies. The user earns points by creating new annotations or by finishing challenges. These points are then used to calculate the user's progression in the application, and if the score obtained is equal to or higher than the goal, the user level goes up, thus giving the user a sense of progress. Information about points and level is shown in the user profile view and in the navbar at the top (Fig. 8).

3.5.3 Rewards

Whenever users complete a challenge, the system rewards them with boosters, points and achievements. Boosters, or point boosters, multiply points to help users to progress. These rewards were motivated to induce desirable behaviours, such as accessing the application regularly to annotate movies, compete with other users to gain a good reputation, towards being the 1st in the global and movie leaderboards.

3.5.4 Leaderboards

Leaderboards or Ranking Tables are an element of gamification related with the needs for esteem and

belonging (Fig.1): a recognition, possibly with a bit of competitiveness, as being first means being best. To unlock that potential, two leaderboards were created: movie and global. In the global one (accessed at movie leaderboard, navbar and profile view), positions are based on total points accumulated, to increase motivation and stimulate competitiveness of frequent or intense annotators at the broader level. The movie leaderboard presents positions in the annotation of each movie. A user may not have a great position in the global leaderboard, but be a big fan and contributor to specific, even niche, movies; and this is a way to promote and recognize that. The best at each movie could even be part of a global board of movies.

3.5.5 Notifications

Having all these gamification elements evolving along time creates the need for a way to convey the state of progress beyond the profile and navbar. Notifications (Fig. 9) were created to keep the user up to date. The notification system informs the user whenever they level up or initialise a booster, when a booster time has finished, an achievement is unlocked, a challenge is completed; and they also receive a daily challenge that they can accept or reject.

Figure 9: Profile view: a) Level; b) Score; c) Notifications; d) Boosters; and, Challenges, Achievements and Global leaderboard tabs.

4 USER EVALUATION

A preliminary user evaluation was conducted to assess perceived usefulness, usability and user experience of Annotation & Gamification features.

4.1 Methodology

We conducted a task-oriented evaluation with semi-structured Interviews and Observation while the users performed the tasks. After explaining the purpose of the evaluation, asking demographic questions and briefing the subjects about the application, the users performed a set of tasks. For each task, we observed and annotated success of completion, errors, hesitations, and their qualitative feedback through comments and suggestions. There was also an evaluation based on USE (Lund, 2001) for each task, where they rated perceived Utility, Satisfaction in user experience and Ease of use, on a 5-point scale.

At the end, users characterized the application with most relevant perceived ergonomic, hedonic and appeal quality aspects, by selecting pre-defined terms (Hassenzahl et al., 2000) that reflect aspects of fun and pleasure, user satisfaction and preferences.

4.2 Participants

This study had 5 participants, 2 male, 3 female, 24-54 years old (M: 36.8; SD: 16.2), 2 with Bachelor of Science (BSc) and two with high school level education, coming from diverse backgrounds (nurse, aeronautical maintenance technician, 2 IT engineers, and hairdresser), all having moderate to high acquaintance with computer applications and having their first contact with this application.

Most of the participants use these devices to watch movies: TV(4), computer(3), mobile phone(3), or tablet(2), and also have access to movies in cinema theaters, smart tvs or streaming platforms. In general, users find interesting to know the genre of a movie, and they are interested in knowing which emotions movies elicit – globally (M: 4.2; SD: 0.8); within a scene (M: 4.2; SD: 0.8), and continuously throughout the movie (M: 3.6; SD: 0.5). Users agreed that they would like to have movies annotated with emotional data, either manually annotated (M: 4.2; SD: 0.4) or automatically detected (M: 4.2; SD: 0.4). Users would be interested in reviewing annotations (M: 4.2; SD: 0.8) and in comparing annotations (M: 4.4; SD: 0.9). None of the users had previously interacted with music, images or movie systems with emotional data. Most users were familiar with gamification elements such as points, level, challenges and achievements.

4.3 Results

Users finished almost all the tasks quickly and without many hesitations, and generally enjoyed the experience with the application. Tasks were organized in 4 groups, and were followed by an overall evaluation.

Annotation: this group included 10 tasks for the annotation view functionalities, specifically annotating movie instants and segments, timeline affective feedback, zoom and delete annotation. Movie Detail:

consisted of 4 tasks, focused on visualizing and understanding the relevance of having emotional information in movies. ReView: included 8 tasks, and assessed the ability to review and compare annotation sessions. Gamification: the last 13 tasks were focused on the evaluation of the gamification elements implemented throughout the system.

Table 1: USE results (1-5 maximum) for the Annotation, Movie detail, ReView, and Gamification tasks. I: Instant, S: Segment, G: Global leaderboard, and M: Movie leaderboard.

Task		U		S		E	
T#	Feature	M	*SD*	M	*SD*	M	*SD*
	Annotation *(mean)*	4.7	*0.5*	4.4	*0.4*	4.5	*0.6*
1.1	Continuous wheel	**5**	*0*	**5**	*0*	**5**	*0*
1.2	Continuous SAM	4.6	*0.5*	3.6	*0.5*	4.4	*0.5*
1.3	Continuous Categorical	**4.8**	*0.4*	**4.8**	*0.4*	4.6	*0.9*
1.4	Timeline	4.2	*0.8*	4.2	*0.8*	4.2	*0.8*
1.5	Timeline zoom	3.6	*0.5*	3.8	*0.4*	4	*0.7*
1.6	Timeline feedback (I)	4.6	*0.5*	**4.8**	*0.4*	**4.8**	*0.4*
1.7	Choose scene	4.6	*0.9*	4	*0.7*	4.4	*0.9*
1.8	Annotation scene	**4.8**	*0.4*	**4.8**	*0.4*	**4.8**	*0.4*
1.9	Timeline feedback (S)	**4.8**	*0.4*	**4.8**	*0.4*	**4.8**	*0.4*
1.10	Delete annotation	4.6	*0.5*	4.2	*0.4*	4.2	*0.8*
	Movie Detail *(mean)*	4.7	*0.5*	4.8	*0.3*	4.7	*0.5*
2.1	Data visualization (I)	4.6	*0.5*	4.6	*0.5*	4.6	*0.5*
2.2	Timeline visualization	4.6	*0.5*	**4.8**	*0.4*	4.4	*0.5*
2.3	Tag Cloud	4.6	*0.5*	**4.8**	*0.4*	**4.8**	*0.4*
2.4	Data visualization (S)	**4.8**	*0.4*	**5**	*0*	**4.8**	*0.4*
	ReView *(mean)*	4.7	*0.5*	4.7	*0.4*	4.7	*0.5*
3.1	Choose single session	**4.8**	*0.4*	**4.8**	*0.4*	**4.8**	*0.4*
3.2	Review Wheel	4.4	*0.5*	**4.8**	*0.4*	**4.8**	*0.4*
3.3	Review SAM	4.6	*0.5*	**5**	*0*	**5**	*0*
3.4	Review Categorical	**4.8**	*0.4*	**4.8**	*0.4*	4.6	*0.9*
3.5	Review all	4.6	*0.5*	4.2	*0.8*	4.4	*0.5*
3.6	Filter review sessions	**4.8**	*0.4*	**4.8**	*0.4*	4.6	*0.5*
3.7	Choose pair session	4.6	*0.5*	4.6	*0.5*	4.6	*0.5*
3.8	Compare two sessions	**4.8**	*0.4*	**4.8**	*0.4*	4.6	*0.5*
	Gamification *(mean)*	4.6	*0.3*	4.6	*0.3*	4.7	*0.3*
4.1	Notifications	**5**	*0*	**5**	*0*	**5**	*0*
4.2	Daily Challenges	**5**	*0*	**5**	*0*	**5**	*0*
4.3	Score and Level	**5**	*0*	**4.8**	*0.4*	**5**	*0*
4.4	Global Leaderboard	**5**	*0*	**4.8**	*0.4*	**5**	*0*
4.5	Sorting System (G)	**4.8**	*0.4*	**4.8**	*0.4*	**4.8**	*0.4*
4.6	Filter System (G)	4.4	*0.5*	4.4	*0.5*	4.6	*0.5*
4.7	Achievements	4.4	*0.5*	4.4	*0.5*	4.6	*0.5*
4.8	Challenges	4.4	*0.5*	4.6	*0.5*	4.6	*0.5*
4.9	Delete Daily Challenge	4.6	*0.5*	4.6	*0.5*	4.6	*0.5*
4.10	Booster	4.4	*0.5*	4.4	*0.5*	4.6	*0.5*
4.11	Movie Leaderboard	4.2	*0.4*	4	*0*	4.6	*0.5*
4.12	Sorting System (M)	4.4	*0.5*	4.6	*0.5*	4.6	*0.5*
4.13	Filter System (M)	4.6	*0.5*	4.4	*0.5*	4.6	*0.5*
	Global Evaluation	4.6	*0.4*	4.6	*0.4*	4.6	*0.4*

Overall evaluation: users classified the applica-tion with most relevant (as many as they found appropriate) perceived quality aspects (23 positive + 23 negative (opposite)), in 3 categories: ergonomic, hedonic and appeal (Hassenzahl et al., 2000). Simple was the most chosen term. Comprehensible, Control-lable were also chosen by more than half of the subjects. The chosen terms are well distributed among the 3 categories. These results confirm and complement the feedback from the other evaluation aspects.

Table 2: Quality terms users chose to describe the evaluated features of the AWESOME application in terms of: Hedonic (H); Ergonomic (E); and Appeal (A).

Terms	Type	#	Terms	Type	#
Simple	E	5	**Motivating**	A	3
Comprehensible	E	4	Clear	E	2
Controllable	E	4	Original	H	2
Trustworthy	E	3	Good	A	2
Impressive	H	3	Inviting	A	2
Innovative	H	3	Sympathetic	A	2
Pleasant	A	3	Exciting	H	1
Attractive	A	3	Desirable	A	1

A summary of the results is presented in tables 1 and 2, and further commented in the conclusions.

5 CONCLUSIONS

This paper presented the background and main features of the Emotional Annotation and the Gamification subsystems of a web application that has been designed and developed to access movies based on emotional impact. The annotations, using different emotional representations, help to enrich the classification of movies based on emotions and their impact on users. This subsystem also allows to collect and present these annotations like in a personal journal, allowing users to see them in their respective representations and they can review and even compare previous annotation sessions, to realise how different the same movie impacted them at different stages in their lives. The gamification subsystem is designed to engage, motivate and reward users in their annotation tasks, and for that, the gamification elements like points, level, challenges, achievements and leaderboards are used. Leaderboards support the needs for esteem and belonging; while the progression with points and levels support Pink's intrinsic motivation of mastery, even recognised at the level of specific movies and their fans; and the challenges can account for the motivation of purpose; all these related with the self-actualization need of Maslow (presented in section 2.4).

A task-oriented user evaluation was carried out, based on semi-structured interviews and observation; each task rated by Utility, Satisfaction and Ease of use, and users commented about current and future developments, in a participatory design perspective. On average, the evaluation score was very high, meaning that the users appreciated and enjoyed the experience of annotating movies, reviewing and comparing annotation sessions, with the gamification elements. In the task to annotate continuously with the self-assessment manikin, users suggested to

replace the 2 dimensions' grids with a matrix that contains all combinations in one selection. In the categorical emotion annotation, some users suggested that the emotions could be represented by emojis or images, as they are quite pervasive nowadays; and they confirmed the need for some adjustments in the timeline zoom, influencing their optimal experience. For continuous annotation they were unanimous about the wheel is the most useful, satisfactory and easy to use, and the SAM the least adequate with 2 values to input. Regarding gamification, users were quite fond of the daily challenges and notifications, then the scores and leaderboards. In particular, someone mentioned getting motivated when accessing the global leaderboard, as an incentive to continue and annotate more movies.

For the future, in the ReView we want to make possible for the users to compare their annotations with other users and even with those of the movie director or other relevant people in this context, like the actors or movie experts. Also, adding annotation by content would be a relevant perspective, allowing to find out, for example, the emotional impact of scenes with screams, I hate/love you declarations or specific music moods. For gamification, an administration profile would enrich and ease creating new achievements and daily challenges.

We believe that these emotional features could be a service for everyone interested in movies and in contributing with their annotations, while increasing their emotional awareness about the movies they watch over time, and keeping the ones they treasure the most.

ACKNOWLEDGEMENTS

This work was partially supported by FCT through funding of the AWESOME project, ref. PTDC/CCI/ 29234/2017, and LASIGE Research Unit, ref. UIDB/00408/2020.

REFERENCES

Ahn, L. V. (2006). Games with a Purpose. IEEE Computer Magazine, 96-98.

Ahn, L. V., Ginosar, S., Kedia, M., & Blum, M. (2007). Improving Image Search with PHETCH. 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry, 25(1), 49–59.

Calvo, R. A., and Peters, D., (2014). *Positive computing: Technology for wellbeing and human potential*. Cambridge, MA: MIT Press.

Chambel, T., & Carvalho, P. (2020). "Memorable and Emotional Media Moments: reminding yourself of the good things!". *Proceedings of VISIGRAPP/HUCAPP 2020*, pp.86-98, Valleta, Malta, Feb 27-29, 2020.

Chitu, A. (2016, Julho 30). Google Image Labeler Is Back. http://googlesystem.blogspot.com/2016/07/google-image-labeler-is-back.html

Choi, S., & Aizaha, K. (2019). Emotype: Expressing emotions by changing typeface in mobile messenger texting. Multimedia Tools and Applications, 14155-14172.

Cowie, R., Douglas-Cowie, E., Savvidou, S., McMahon, E., Sawei, M., & Schröder, M. (2000). "Feeltrace": An instrument for recording perceived emotion in real time. ISCA Workshop on Speech {&} Emotion, 19-24.

Csikszentmihalyi, M., Abuhamdeh, S., & Nakamura, J. (2014). Flow. In *Flow and the foundations of positive psychology* (pp. 227-238). Springer, Dordrecht.

Deterding, S., Dixon, D., Nacke, L.E., O'Hara, K., Sicart, M. (2011). Gamification: Using Game Design Elements in Non-Gaming Contexts. In Proc. of CHI EA'11. Workshop. ACM. 2425-2428.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, 169-200.

Garrity, A., & Schumer, L. (2019). *What Is a Bullet Journal? Everything You Need to Know Before You BuJo*. Good House Keeping. https://www.goodhousekeeping.com/life/a25940356/what-is-a-bullet-journal/

Girard, J. M., & Wright, A. G. (2018). DARMA: Software for dual axis rating and media annotation. Behavior Research Methods, 902-909.

Hantke, S., Eyben, F., Appel, T., & Schuller, B. (2015). IHEARu-PLAY: Introducing a game for crowdsourced data collection for affective computing. 2015 International Conference on Affective Computing and Intelligent Interaction, ACII 2015, 891-897.

Hassenzahl, M., Platz, A., Burmester, M, Lehner, K. (2000) Hedonic and Ergonomic Quality Aspects Determine a Software's Appeal. *ACM CHI 2000*. The Hague, Amsterdam, pp.201-208.

Hildebrand, M., Brinkerink, M., Gligorov, R., Van Steenbergen, M., Huijkman, J., & Oomen, J. (2013). Waisda? Video labeling game. *MM 2013 - Proceedings of the 2013* ACM Multimedia Conference, 823-826.

Horner, H. (2018). The Psychology of Video: Why Video Makes People More Likely to Buy, *Sproutvideo*. https://sproutvideo.com/blog/psychology-why-video-makes-people-more-likely-buy.html

Humane Engineering Ltd., Cove: music for mental health, 2015-19. https://apps.apple.com/app/cove-the-musical-journal/id1020256581

Huotari, K., & Hamari, J. (2012). Defining gamification - A service marketing perspective. *Proceedings of the 16th International Academic MindTrek Conference*: "Envisioning Future Media Environments", MindTrek 2012, 17 - 22.

Khaled, R. (2011). It's Not Just Whether You Win or Lose: Thoughts on Gamification and Culture. In Proc. of Gamification Workshop at ACM CHI'11.

Law, E., & Ahn, L. V. (2009). Input-agreement: A new mechanism for collecting data using human computation games. Conference on Human Factors in Computing Systems - Proceedings, 1197-1206.

Law, E. L., Ahn, L. v., Dannenberg, R. B., & Crawford, M. (2007). TagATune: A Game for Music and Sound Annotation. Proceedings of the 8th International Conference on Music Information Retrieval, 23-27.

Lund, A. M. (2001). Measuring usability with the USE questionnaire. Usability and User Experience, 8(2).

Maslow, A. H. (1943). A theory of human motivation. *Psychological review*, 50(4), 370.

Marczewski, A. (7 of May of 2014). GAME: A design process framework. Obtido de Gamified UK: https://www.gamified.uk/2014/05/07/game-design-process-framework/

McLeod, S. (2020). *Maslow's Hierarchy Needs*, Simply Psychology.https://www.simplypsychology.org/maslow.html

Mora, A., Riera, D., Gonzalez, C., & Arnedo-Moreno, J. (2015). A Literature Review of Gamification Design Frameworks. VS-Games 2015 - 7th International Conference on Games and Virtual Worlds for Serious Applications, 100-117.

Nave, C., Correia, N., & Romão, T. (2016). Exploring Emotions through Painting, Photography and Expressive Writing: an Early Experimental User Study. In Proc. of the 13th International Conf. on Advances in Computer Entertainment Technology (pp. 1-8).

Oliveira, E., Martins, P., & Chambel, T. (2013), "Accessing Movies Based on Emotional Impact", Special Issue on "Social Recommendation and Delivery Systems for Video and TV Content", *ACM/Springer Multimedia Systems Journal*, ISSN: 0942-4962, 19(6), 559-576.

Picard, R. W. (2000). Affective computing. MIT press.

Pink, D. H. (2011). *Drive: The surprising truth about what motivates us*. Penguin.

Plaisent, M., Tomiuk, D., Pérez, L., Mokeddem, A., & Bernard, P. (2019). *Serious Games for Learning with Digital Technologies*. Springer, Singapore.

Plutchik, R. (1980). Emotion: A Psych evolutionary Synthesis. Harper and Row.

Russell, J. A. (1980). A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161–1178. https://doi.org/10.1037/h0077714.

Sherer, K. R. (2005). What are emotions? And how can they measured? *Social Science Information, 44*(4), 695-729.

Souza, M. R., Moreira, R. T., & Figueiredo, E. (2019). Playing the Project: Incorporating Gamification into Project-based Approaches for SW Eng. Education.

Sreeja, P. S., & Mahalakshmi, G. S. (2017). Emotion Models: A Review. International Journal of Control Theory and Applications, 651-657.

Viana, P., & Pinto, J. P. (2017). A collaborative approach for semantic time-based video annotation using gamification. Human-centric Computing and Information Sciences, 1-21.

Wiegand, T., & Stieglitz, S. (2014). Serious fun-effects of gamification on knowledge exchange in enterprises. Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI), 321-332.

Wu, M. (2011) *Gamification 101: The Psychology of Motivation*, Khoros.https://community.khoros.com/t5/Khoros-Communities-Blog/Gamification-101-The-Psychology-of-Motivation/ba-p/21864

Zhang, T., El Ali, A., Wang, C., Hanjalic, A., & Cesar, P. (2020). RCEA: Real-time, Continuous Emotion Annotation for Collecting Precise Mobile Video Ground Truth Labels. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).