Characterizing Player's Experience From Physiological Signals Using Fuzzy Decision Trees

Florent Levillain, Joseph Onderi Orero, Maria Rifqi and Bernadette Bouchon-Meunier

***Abstract*— In the recent years video games have enjoyed a dramatic increase in popularity, the growing market being echoed by a genuine interest in the academic field. With this flourishing technological and theoretical efforts, there is need to develop new evaluative methodologies for acknowledging the various aspects of the player's subjective experience, and especially the emotional aspect. In this study, we addressed the possibility of developing a model for assessing the player's enjoyment (amusement) with respect to challenge in an action game. Our aim was to explore the viability of a generic model for assessing emotional experience during gameplay from physiological signals. In particular, we propose an approach to characterize the player's subjective experience in different psychological levels of enjoyment from physiological signals using fuzzy decision trees.**

I. Introduction

Fun is a crucial component of video games. Video games are purposely designed to elicit positive experiences, where every obstacle standing on the player's path should be an excuse for entertainment. Yet, when it comes to defining the factors determining such an experience, quantifying fun remains a complex task. In the domain of game design, empirical methods still rule the process of making a game enjoyable, and objective and systematic methods are still lagging behind the sheer interest for theoretical approaches in game design. In the recent past, research has focused on machine learning approaches, with the goal of modeling emotional experiences related to gameplay. In this direction, considerable progress has been made in using physiological signals as a source of input [32], [22], [7], [41].

However, despite these promising scientific advances, physiological computing still face a number of obstacles [12]. Fundamental issues such as the generality and standardization of the methodologies developed [1], are yet to be fully addressed. The approaches tend to be too specific and dependent on the laboratory experiments, making it difficult to compare results and validate their applicability in real-time applications. Indeed, in order to realize the full potential of affective computing, more emphasis should be put in developing generic user models that represent the players [15]. The aim of this work was to address some of these fundamental challenges.

Florent Levillain is with the Laboratoire Cognitions Humaine et Artificielle (CHART), Université Paris 8, France (email: flevillain@mac.com).

Joseph Onderi Orero, Maria Rifqi and Bernadette Bouchon-Meunier are with the Laboratoire d'Informatique de Paris 6 (LIP6), Université Pierre et Marie Curie, France (email:{joseph.orero,maria.rifqi, bernadette.bouchon-meunier}@lip6.fr).

First, it is necessary to express in fuzzy terms the mapping of affective markers from physiological data. It can be argued that, changes from one emotional state to the next is gradual rather than abrupt and that we need to take into account the overlapping of class boundaries [14]. Moreover, the physiological data from sensors is itself imperfect, such that it is difficult to express the results in crisp terms [2]. Fuzzy set theory based models seem more applicable to represent these continuous transitions, uncertainties and imperfections. As a result, fuzzy set theory based approaches have been proposed with promising prospects to assess player's satisfaction [22], [32], [39]. Nevertheless there is need to advance further in this direction, especially by exploring methods able to extract relations that define the optimal combination of measures. Despite their many advantages in real-time applications, physiological measurements do not provide a lateral, isomorphic representation of the emotion or intention [12]. Consequently, in this work, we employ fuzzy decision trees to extract psychophysiological relations.

Secondly, to guarantee the viability of the developed models in real-time application, the experimental setup should be as close to normal human situations as possible. In order to ensure a natural sense of immersion in the players, we put some efforts in creating an adequate experimental situation. We recruited our participants with no other incentive than to have fun playing the game. To respect their spontaneous pace, we placed no constraints on the way the game has to be played. We chose a popular game, well known for its smooth control system and its sense of balance and immersion but selected episodes in the game with a clear contrast in terms of difficulty, although each of them was worth playing. Finally, to gather different layers of players'satisfaction, we used two types of questionnaires: one immediately after each sequence, to probe the most recent memories of the feelings elicited by the game; another one, more retrospective, at the end of the session, to promote a comparison between the different sequences.

Overall, our main goal was to distinguish typical physiological signatures associated with various gaming experiences. In order to correlate these physiological measures with subjective evaluations in different psychological levels of enjoyment, we recorded physiological signals with the hope of finding features able to distinguish between the levels of engagement elicited by these sequences.

The rest of the paper is organized as follows: in Section II we outline how to assess player's experiences. In Section III we give justification of machine learning approaches. We then describe the experimental setup in Section IV before giving our results in Section V. Finally, we give conclusions and future perspectives in Section VI.

II. How To Assess Player Experiences

A. Objective Measures

Until now, modeling of the player's emotional experience during gameplay has mainly relied on traditional methods through subjective self-reports such as questionnaires, interviews and focus groups. Although these subjective self-report methods are used virtually in all fields to give a global view of the user experiences on specific aspects of interest, they have limitations. One of them being that they only generate data at the moment the question is asked, and not through a continuous process. Secondly, it is difficult for participants to self report their behaviors during game situations [28]. Since the main goal of video games is to entertain through a continuous renewal of the user's interest, controlling the emergence of certain affective states is crucial in achieving a truly immersive experience. Therefore, in evaluating video games, it would be more appropriate to use more objective measures that can assess continuously the emotional responses in relation to variations of scenarios and tasks at hand in the game.

Among a vast range of possible ways to continuously assess a user's emotional responses such as facial gesture or voice recognition through video and audio recording, physiological measures stand out. Although physiological recordings have problems of their own (they produce noisy data, they are not yet fully wearable and require a certain immobility from the user), they grant an access to non conscious and non reportable processes [4] and may to a certain extent be unobtrusively monitored [35]. Since video games tend to promote a natural sense of immersion, physiological recordings seem appropriate when it comes to measuring how much cognitive effort, or active coping is involved in a particular task. Thus in this work, we focused on the use physiological signals as a more objective measure to continuously assess the players' emotional experience.

B. Modeling the Player's Experience

There is much debate concerning how to conceptualize emotions, whether they must be considered as static, biologically-rooted states [11] or as dynamical, boundaries-free states [27]. The dimensional theory of emotions holds that all emotions can be resumed as coordinates of valence and arousal [18], [34]. For instance, in the domain of game studies ([22], [5] and many others), the player's experience is defined as a compound of a certain amount of arousal and a certain degree on a valence scale.

But when trying to model the player's subjective experience, an approach based only on valence and arousal might come short. Specifically, this approach does not take into account the possibility that the player's satisfaction comes from two different sources: (i) the (meta)cognitive assessment of the challenge at stake, reflecting the recognition of the game designer's intentions to manipulate the affective/immersive component of the game, (ii) the affective evaluation of the pleasure gained from experiencing a certain amount of challenge. In other words, a player may recognize the intention of the game to vary the level of excitement, although he may not enjoy such a variation. This is coherent with a dual-system of evaluation based on distinct anatomical pathways, a cognitive pathway and an emotional pathway [19].

In this respect, considering the player's appraisal of the challenge at stake might be a more promising approach. In this domain, one key approach concerns the theory of flow [9]. This theory states that in order to elicit positive emotions in the player, one should be certain that the player is maintained in a narrow channel (the flow channel) where he/she is not overcome by difficulty, although at the same time challenge should not fail to engage the player. The limits of the state of flow are thus boredom in the one hand, when the player feels insufficiently engaged, anxiety in the other hand, when challenge exceeds the player's skills. To put it in simple words, a game should be neither too hard nor too easy [13], which is represented by a recent tendency in game design to address several levels of skills by promoting multi-layered games. In order to satisfy the player, therefore to maintain a state of flow as frequently as possible, it is thus extremely important to learn to recognize, with the help of objective indicators, what is a state of optimal satisfaction in his/her experience.

C. Application of Gameplay Experience Modeling

The approach of modeling player's experience with respect to appraisal of challenge, has been widely explored, especially in the recent past ([7], [40], [32] among others). Although the methodologies may seem similar, there seems to be two research concerns in this direction. One approach concerns the qualitative evaluation of the game to ensure that the final product gives the desired experience. In this case, the interest is more on determining the nature of combination of factors that make a computer game fun [21]. A game should be designed to combine these aspects such that it leads to the best experience. For instance, Yannakakis and his team [39], [40], [41] have carried out extensive work developing models able to recognize games designed to give *more fun* with respect to Malone's factors [21].

On the contrary, a second approach seeks to hold certain factors constant, while varying certain aspects of interest. Unlike the former approach, this approach is more concerned by the implicit manipulation of the player's experience during interaction. When it comes to the evaluation of challenge, triggering an optimal experience implies the player's ability to handle the task at hand while being actively engaged. The assumption is that for well-designed game sequences, the appreciation will vary depending on the mastery of the skills required. For instance, Rani *et. al.* [33], [32] induced different levels of anxiety by varying the challenge. In a similar way, Chanel *et. al.* [7] tested the hypothesis that the experience in a game level depends on the player's mastering of the skills required in that specific level. In this direction, a major concern is to enable the possibility for games to adjust online the level of difficulty based on the player's skills.

Although the two approaches are closely interrelated, our focus in this work was more on the second approach. We propose to advance further in this direction by proposing a machine learning approach that could be more applicable in real-time applications. The aim was to extract information for developing a generic emotionally adaptive control [15], that induces a given experience on the player through challenge variation. As outlined by Fairclough's [12] four zones of distress and engagement, the game should maintain a continuous loop between the stretch zone (when engagement and distress are both high) and comfort zone (when the user is comfortable with the level of demand yet remains motivated by the task at hand). The corrective mechanisms of the emotionally adaptive control will depend on whether the player is experiencing:

i) Low Distress
   - Game Engagement (Comfort zone)
   - Game Disengagement *

ii) High Distress
   - Game Engagement (stretch zone)
   - Game Disengagement *

Therefore, our experimental setup was geared towards extracting the characteristics of the physiological features that can be used to discriminate these experiences for an adaptive control to provide appropriate corrective mechanism.

III. Machine Learning

A. Classification Methods

A wide range of methods have been proposed for emotion recognition from physiological signals such as linear discriminant analysis, k-nearest-neighbor (KNN), neural network and decision trees [30], [25], [37], [32], [16]. Although the results from these methods seems comparable (see in [37]), decision trees appear to be better suited for this kind of problem. A major advantage of decision trees lies in the fact that they induce explicitly defined rules used in classification. Our interest is not only to achieve high rates of classification but also to determine the relationship between the physiological signals attributes and the emotional states. The success of affective computing [29] depends on establishing the optimal combination measures and features to discriminate emotional categories. Unlike other classifiers, decision trees not only perform classification but also evaluate attributes by selecting the best attribute that discriminate the classes in each node of the tree. Indeed, with a reasonable tree pruning and sample size, decision trees give characterization of the training set indicating how attribute values differ between different classes.

B. Fuzzy Decision Trees

As we have already pointed out, it is preferable to use an approach based on fuzzy sets theory. In fuzzy sets theory [42], a fuzzy set is represented by a membership function, $\mu_A : A \rightarrow [0,1]$, indicating the degree to which an element belongs to a given set A. This is a contrast to $\{0,1\}$ in a crisp set, in which an element can only belong to a given set A (membership value of 1) or not (membership value of 0). It is interesting for this kind of recognition to express our output in a gradual scale $[0,1]$ especially in order to continuously assess the emotional change during gameplay. In this respect, fuzzy expert systems designed with rules based on psychophysiological literature [22], [32] and learning from fuzzy neural networks [39], have been explored.

In this work, we consider a machine learning by means of fuzzy decision trees. Fuzzy decision trees automatically construct from the data a set of fuzzy rules which is knowledge base for a fuzzy expert system. This approach is an automatic method to build fuzzy partitions from attributes to avoid prior definition of fuzzy values and enables us to test and compare them in the process of classification. In addition, like classical decision trees, they represent induced knowledge in a very expressive way in which the path of a decision tree is equivalent to an *IF ... THEN ...* rule. This is a contrast to black box methods such as neural networks in which the model is represented by values of a network weights. Moreover, in fuzzy decision trees, as changes from one rule to another is gradual with fuzzy values $[0,1]$ instead of crisp values$\{0,1\}$ in classical trees, they have proved to provide better classification rate [23].

The objective of this work was to explore automatic generation of psychophysiological relations based on the players' physiological data. We used Marsala's Salammbô Fuzzy Decision Tree [23] and compared our results with Quilan's C4.5 decision trees [31] and KNN [8].

IV. The Experimental Setup and Setting

A. Participants and Setting

The experiment was conducted at the LUTIN (Laboratoire des Usages en Technologies d'Information Numérique), Paris, France. Participants in the experiment were recruited from visitors of the nearby museum, Cité des Sciences et de l'Industrie. They were all aged between 15 to 39 and no specific expertise in the field of video games was required, although we selected participants able to manipulate a gamepad and to orient themselves in a virtual environment. We tested a total of 39 participants. However, due to failures in the physiological recording, we kept data from 25 participants for whom all the sensors worked.

We tested a game belonging to a popular genre in the game industry, Halo3, which is a First-Person Shooter (FPS). This game is one of the best genre that promotes a sense of immersion in that it propels the player at the heart of action through a first-person perspective. Halo3^{TM} was played on a Microsoft Xbox 360^{TM} on a 32-inch LCD television. A camera captured the TV screen. Participants were seated at approximatively one meter from the screen, they were explicitly told not to move and keep the game pad onto their laps in order to avoid any muscular artefact in the physiological recording.

B. Physiological Signals and Features

In order to discriminate emotions from physiological signals, a wide range of measures have been proposed such as electromyography measuring facial muscle tension, the blood volume pulse, the skin conductance, respiration rate and measures related to the brain activity. In this work, based on the previous literature, we identified a subset of these measures that can be used almost non-intrusively while yielding optimal results. We chose to collect galvanic skin response (GSR), heart rate (HR) and respiration rate (RR) data during gameplay. GSR which is a measure of the conductivity of the skin is considered as an effective correlate to arousal [18], [3], [10] and has been extensively used in the domain of affective computing [35], [38], [22]. On the other hand, heart rate (HR) and blood pressure may also give an indication about stress-related activities with heart rate accelerations mediated by the sympathetic nervous system [20], [24]. But as a result of a dual innervation of the heart by both the sympathetic and the parasympathetic nervous systems, HR could also index moments of attentional surge. For instance, increased cardiac parasympathetic activity causes HR to decelerate when attention is paid to an external (e.g., media) stimulus [17][36].

To collect the physiological measures we used the Biopac $MP35$ acquisition unit and the software BSLPro to visualize the data. We collected heart rate (HR) through a measure of cardiovascular activity by measuring the electrocardiography (ECG) through a Einthoven derivation II placing pregelled surface electrodes on the ankles and on the wrist. We recorded GSR using surface electrodes attached with Velcro^{TM} Straps that were placed on two fingers of the left hand. The fingers wearing the electrodes remained wedged under the gamepad. We recorded the respiration rate (RR) with a stretch sensor positioned around the thorax. ECG, GSR and RR data were collected at 200Hz. As noisy ECG data may produce failures in computing the HR, we inspected the HR data and corrected manually every erroneous samples. The same method was applied to the RR.

In respect to physiological feature extraction, for each signal, we chose to calculate the features shown in Table I that we belief are very relevant based on results from past research and our earlier preliminary work [26].

TABLE I
FEATURES FROM PHYSIOLOGICAL SIGNALS (GSR,HR,RESP)

Feature	Total
Maximum value of raw signal	3
Minimum value of raw signal	3
Mean value of raw signal	3
Standard deviation of raw signal	3
Mean of absolute first derivative of raw signal	3
Maximum gradient of the raw signal	3
Power Spectrum Density 0-0.8 frequency range ($\Delta 0.2$)	12
Total	30

C. Game Sequences

As a baseline, we used a resting period during which participants watched the introductory screen from the game. This screen is appropriate to elicit a relaxing state in that it depicts a contemplative scene with slowly moving objects accompanied by a soft soundtrack. Participants played successively to four short sequences, each of them followed immediately by a questionnaire and a two-minutes resting period. The game sequences were as follows:

i) Sequence 1: The game session always started with an introductory sequence corresponding to the first minutes of the game. In this sequence, the transition from exploration to combat is smooth, and specifically designed not the challenge excessively the player enjoying the game for the first time. After having played Sequence 1, participants were asked to complete the three following sequences. The order of presentation of these sequences was counterbalanced.
ii) Sequence 2: the player is driving a powerful tank, he is heavily protected and benefits from a highly effective arsenal;
iii) Sequence 3: the player is equipped with a sniping riffle, and is allowed to shoot enemies at a distance. The player is in a tactical advantage, as he stands in a upper position;
iv) Sequence 4: the player is confronted to the highest difficulty level of the game. The player is equipped with a very basic arsenal, whereas ennemies are well armed, ferocious and very resistant.

D. Self-Assessment Reports

At the end of each game episode, we asked the participant to rate both her evaluation of the level of certain psychological parameters, as well as the pleasure gained from these parameters on a six-point ranking scale such as:

i) *How much concentration is required in this sequence?*
ii) *Did you enjoy that the sequence requires this particular amount of concentration?*
iii) *How arousing is this sequence?*
iv) *Did you enjoy that the sequence elicits this particular amount of arousal? ....*

Then, at the end of the experiment (after the participant had completed the four sequences), we asked him to answer four questions where a comparison between the four sequences was proposed:

i) *Which sequence is the most amusing?*
ii) *Which sequence is the least amusing?*
iii) *Which sequence is the most challenging?*
iv) *Which sequence is the least challenging?*

V. RESULTS

A. Learning Set Construction

To have a more objective correlation of the subjective experiences and the physiological measures during classification, we used the evaluation at the end of the experiment. As shown in Figure 1 and Figure 2, we found that the sequences considered as the most/the least challenging were not the ones considered as the most amusing. Specifically, we can see that the Sequence 4, which were clearly used as a representative of a very difficult game episode, was chosen by almost every participant as the one considered the most challenging. On the other hand, few participants considered this sequence as the most amusing. This probably reflects the fact that participants felt their skills exceeded in this episode, with a feeling of frustration as a consequence. On the opposite side, the Sequence 1, which was the introduction of the game, was clearly picked up as the least challenging sequence, no participant chose it as the most amusing. In this case, the lack of challenge is certainly the factor determining a weak sense of fun. These two results confirm the fact that in order to optimize the sense of fun, the player must be maintained in a state of flow, between boredom and anxiety, with the Sequence 2 and Sequence 3 standing here in the closest position to this state of flow.

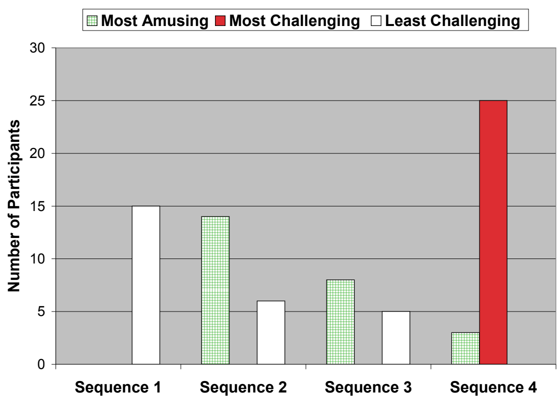

Fig. 1. Most Amusing vs Challenge

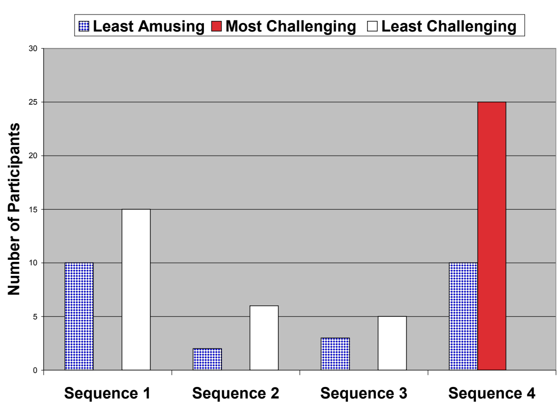

Fig. 2. Least Amusing vs Challenge

First, since there was a variation of the length of the episodes from participant to participant and to minimize the effect of the transition periods, we used only the physiological recordings of the last two minutes of the game sequence. Then we subdivided the signals into 10 seconds (2000 data points) segments. As attributes for the classifiers, we calculated the features explained in Section IV. Secondly, to account for variations between participants, we calculated each participant's attribute normalized value, nA_i, from the row value, A_i, using the attribute's standard deviation, A_{sdv}, and its mean, A_{mean} as shown in Equation 1.

$$nA_i = \frac{A_i - A_{mean}}{A_{sdv}} \tag{1}$$

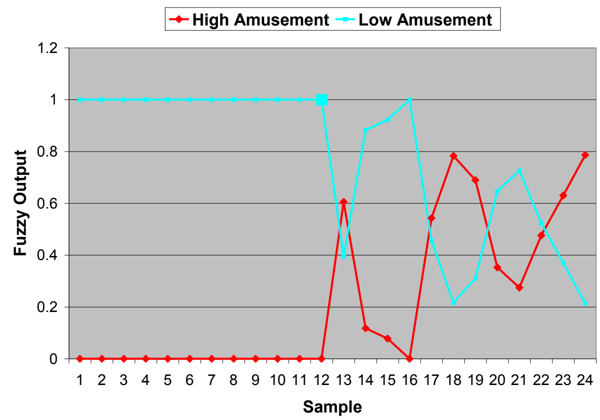

Fig. 3. Example of Output from Fuzzy Decision Tree

The output from the classifiers was in linguistic variables for each of the dimensions. The output from the fuzzy decision tree was in continuous values $[0, 1]$ indicating the membership grades of a given sample to a particular linguistic variable. Figure 3 shows an example of fuzzy decision tree output from one of the participants. Samples 1 to 12 were samples from the sequence evaluated as the least amusing sequence while the rest were from the sequence evaluated as the most amusing. As it can be seen, although some samples from the least challenging sequence were classified as high and vise versa, we have more information as regards to the gradual change from one point in time to the next. This kind of information is critical towards building expert systems. We defuzzified the output into values $\{0, 1\}$ and compared the results with other classifiers.

B. Results of Classification

Our main objective was to extract physiological features that characterize player's level of enjoyment (amusement). In order to have two clearly distinct categories of game sequences, we contrasted between the game sequence identified as the most amusing against the one identified by the player as the least amusing. We obtained the results shown in Table II. Figure 4 shows a sample of decision tree for discriminating most and least amusing game sequences. In particular, when GSR mean was greater than 0.22, the majority of the sequences are of low amusement.

Alternatively, as already discussed we can consider the amusement with respect to challenge. As we already noted, we can distinguish two possible undesirable player

TABLE II

MOST/LEAST AMUSING GAME SEQUENCES CLASSIFICATION RESULTS

	Salammbô FDT	Quilan C4.5 DT	KNN(k=10)
Least Amusing	80.40%	78.80%	77.50%
Most Amusing	71.30%	70.00%	76.30%
Overall	75.85%	74.40%	76.90%

Low/High Amusement Decision Tree

GSR mean <= 0.22266

| HR min <= 67.90785

| | RR derivative<= 3.02265:Low Amusement (21):High Amusement (6)

| | RR derivative > 3.02265

| | | RR min <= 16.44860: Low Amusement (4)

| | | RR min > 16.44860: High Amusement (17): Low Amusement (2)

| HR min > 67.90785: High Amusement (89): Low Amusement (20)

GSR mean > 0.22266

| GSR min <= 6.33326: Low Amusement (78): High Amusement (12)

| GSR min > 6.33326

| | GSR min <= 8.76122: High Amusement (14)

| | GSR min > 8.76122: Low Amusement (19): 1.00 (6)

Fig. 4. Sample Extract of High/Low Amusement Decision Tree

experiences (a) moments disengagement due to low distress , when the player is likely to feel insufficiently challenged by the task (too low distress disengagement) and (b) those moments when the player is overstretched beyond his/her ability leading to anxiety (high distress disengagement). As our interest is in developing emotionally adaptive system, it is important to identify the physiological features that characterize them as they require different adaptive mechanisms. First, as shown from Figure 1, Sequence 1 which was the least challenging game sequence, was not identified by any participant as the most amusing. Therefore, to extract physiological features that characterize a game sequence of low distress disengagement (low challenge), we contrasted Sequence 1 and the sequence identified by the participant as the most amusing and found the results shown in Table III. We pruned a sample decision tree to produce the following two rules:

GSRmin$<= -0.33239 : TooLowDistress(104) : MostAmusement(19)$

GSRmin$> -0.33239 : MostAmusement(125) : TooLowDistress(40)$

The GSR signal can thus be successfully used to detect periods when the player is insufficiently challenged (too low distress disengagement).

TABLE III

LOW DISTRESS DISENGAGEMENT/MOST AMUSEMENT SEQUENCES CLASSIFICATION RESULTS

	Salammbô FDT	Quilan C4.5 DT	KNN(k=10)
Too Low Distress	77.40%	79.90%	80.60%
Most Amusing	86.30%	82.30%	78.50%
Overall	81.85%	81.10%	79.50%

Secondly, in order to ascertain the physiological features that characterize moments when the player feels overloaded by the task (overload disengagement), we contrasted the sequences judged as the most challenging with the one identified as the most amusing. As shown in Figure 1, Sequence 4 is clearly judged as the most challenging, but 22 participants out of 25 did not find it the most amusing. Using data from these 22 participants, we contrasted between sequences identified by the player as most amusing against Sequence 4 (the most challenging sequence). We obtained the results shown in Table IV. Globally, the generated decision trees revealed that, the GSR (min), HR (min amplitude value, maximum value of the first derivative) and RR (maximum value of the first derivative and power spectrum density), were the most relevant features.

TABLE IV

OVERLOAD DISENGAGEMENT/MOST AMUSEMENT SEQUENCES CLASSIFICATION RESULTS

	Salammbô FDT	Quilan C4.5 DT	KNN(k=10)
Overload	92.59%	78.20%	55.60%
Most Amusement	62.37%	65.50%	75.90%
Mean	69.17%	72.10%	65.40%

VI. CONCLUSIONS AND FUTURE PERSPECTIVES

In this study, we set up an experiment to enable us model the player's experiences. We used fuzzy decision trees to automatically characterize the behavior of physiological signals with respect to players' evaluation of challenge in a game episode. We managed to identify with considerable success amusement level in respect to variation of challenge at stake. Our results thus show that it is possible to gain information from physiological signals considering the optimal state of satisfaction of a player. Flow, in terms of physiological activity, is reflected by variations in GSR and other range of physiological activations that may be considered as the sign of a deeper immersion in the game.

Eliciting a state of higher arousal by increasing the challenge faced by the player is one of the component necessary to get a positive reaction. However, this component, when not coupled to a possible immersion component, may impede the sense of positive involvement in the game. In our experiment, this seems to happen especially when players face the highest level of difficulty of the game. Therefore, flow might be characterized as a sense of high physiological arousal coupled with a (possibly more cognitive) feeling of adequation between the level of difficulty and the skills at hand.

However, much is still to be done before getting access to the structure of the player's emotional processes. Our results further confirm the difficulty in performing machine learning due to inter individual variations of physiological signals. Physiological signals seem to vary considerably from participant to participant. Indeed, although we attempted to minimize these variations by normalizing the features for each participant, it may not have been successful due to enormous variations between individuals' physiological data. We thus need to consider better algorithms to tackle this problem.

Altogether, the road map for the forthcoming investigation of affective states in video game will get through a clear definition of the most relevant dimensions to account for the emotional response we target, as well as a thorough examination of other machine learning approaches. In this work, we introduced the aspect of correlating objective with subjective measures. Some more work is needed to understand how to combine subjective measures with multiple objective measures. This way, we hope for a truly systematic affective recognition procedure to be incorporated to the games evaluation routines.

References

[1] J. Arroyo-Palacios and D. M. Romano. Towards a standardization in the use of physiological signals for affective recognition systems. In *Measuring Behavior*, 2008.

[2] B. Bouchon-Meunier. *Aggregation and Fusion of Imperfect Information*. Physica-Verlag, Spring-Verlag Company, 1998.

[3] M. Bradley, M. K. Greenwald, and A.O. Hamm. Affective picture processing. In N. Birbaumer and A. Ohman, editors, *The Structure of Emotion*, pages 48–65. Hogrefe and Huber Publishers, Toronto, 1993.

[4] J. T. Cacioppo, L. G. Tassinary, and G.G. Bernston. Psychophysiological science: Interdisciplinary approaches to classic questions about the mind. In G.G. Bernston J.T. Cacioppo, L.G. Tassinary, editor, *Handbook of Psychophysiology*. Cambridge University Press, 2007.

[5] G. Chanel, K. Ansari-Asl, and T. Pun. Valence-arousal evaluation using physiological signals in an emotion recall paradigm. In *Ieee Smc And International Conference On Systems, Man And Cybernetics, Smart Cooperative Systems And Cybernetics: Advancing Knowledge And Securityfor Humanity*, 2007.

[6] G. Chanel, J. J. M. Kierkels, M. Soleymani, and T. Pun. Short-term emotion assessment in a recall paradigm. *Int ernational Journal of Human-Computer Studies*, 67:607–627, 2009.

[7] G. Chanel, C. Rebetez, M. Bétrancourt, and T. Pun. Boredom, engagement and anxiety as indicators for adaptation to difficulty in games. In *The 12th International Conference on Entertainment and Media in The Ubiquitous Era*, 2008.

[8] T.M. Cover and P.E. Hart. Nearest neighbor pattern classficiation. *IEEE Transactions Information Theory*, 13:21–27, 1967.

[9] M. Csikszentmihalyi. *Harper and Row Flow: The psychology of optimal experience*. Harper & Row New York, 1990.

[10] B.H. Detenber, R.F. Simons, and G.G. Bennett. Roll' em! the effects of picture motion on emotional responses. *Journal of Broadcasting and Electronic Media.*, 21:112–126, 1998.

[11] P. Ekman, R.W. Levenson, and W.V. Friesen. Autonomic nervous system activity distinguishes among emotions. *Science*, 221:1208–1210, 1983.

[12] S.H. Fairclough. Fundamentals of physiological computing. *Interacting With Computers*, 21:133–145., 2009.

[13] T. Fullerton, C. Swain, and S. Hoffman. *Game Design Workshop: Designing, Prototyping, and Playtesting Games.* San Francisco, 2004.

[14] P. Goulev. *An Investigation into the Use of AffectiveWare in Interactive Computer Applications*. PhD thesis, Imperial College London, 2005.

[15] E. Hudlicka. Affective game engines: Motivation and requirements. In *International Conference On The Foundations Of Digital Games*, 2009.

[16] J. Kim and E. André. Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis And Machine Intelligence*, 30(12):2067–2083, 2008.

[17] J.L. Lacey and B.C. Lacey. *Some autonomic-central nervous system interrelationships*. New York: Academic Press, 1970.

[18] P. Lang. The emotion probe studies of motivation and attention. *American Psychologist*, 50(5):372–385, 1995.

[19] J. Ledoux. *The Emotional Brain: The Emotional Underpinnings of Emotional Life*. Simon and Schuster, 1998.

[20] M. M. Llabre, S.B. Spitzer, P.G. Saab, G.H. Ironson, and N. Schneiderman. The reliability and specificity of delta versus residualized change as measures of cardiovascular reactivity to behavioral challenges. *Psychophysiology*, 28:701–711, 1991.

[21] T. W. Malone. What makes things fun to learn? heuristics for designing instructional computer games. In *the 3rd ACM SIGSMALL symposium and the first SIGPC symposium on Small systems*, 1980.

[22] R. L. Mandryk and M.S. Atkins. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. *International Journal of Human-Computer Studies*, 65(4):329–347, 2007.

[23] C. Marsala. Fuzzy decision trees to help flexible querying. *KYBERNETICA*, 36:689–705, 2000.

[24] J.K. Murphy, C.M. Stoney, B.S. Alpert, and S.S. Walker. Gender and ethnicity in childrens cardiovascular reactivity : 7 years of study. *Health Psychology.*, 14:48–55, 1995.

[25] F. Nasoz, K. Alvarez, C.L. Lisetti, and C.L.N. Finkelstein. Emotion recognition from physiological signals for presence technologies. *International Journal Of Cognition, Technology And Work*, 6(1), 2003.

[26] J. O. Orero, F. Levillain, M. Damez-Fontaine, M. Rifqi, and B. Bouchon-Meunier. Assessing gameplay emotions from physiological signals : A fuzzy decision trees based model. In *INTERNATIONAL CONFERENCE ON KANSEI ENGINEERING AND EMOTION RESEARCH*, 2010.

[27] A. Ortony and T. J. Turner. Whats basic about basic emotions? *Psychological Review*, 97:315–331, 1990.

[28] R. J. Pagulayan, K. Keeker, D. Wixon, R. Romero, and T. Fuller. User-centered design in games. In *Handbook for human-computer interaction in interactive systems*, pages 883–906. Lawrence Erlbaum Associates, Inc. Mahwah, NJ, 2002.

[29] R. W. Picard. *Affective Computing*. The MIT Press, Cambridge, Massachusetts, London, England, 1997.

[30] R.W. Picard, E. Vyzas, and J. Healey. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 23:1175–1191, 2001.

[31] R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo, CA, 1993.

[32] P. Rani, N. Sarkar, and J. Adams. Anxiety-based affective communication for implicit human machine interaction. *Advanced Engineering Informatics*, 21:323–334, 2007.

[33] P. Rani, N. Sarkar, and C. Liu. Maintaining optimal challenge in computer games through real-time physiological feedback. In *11th HCI International, (Las Vegas, USA), Lawrence Erlbaum Associates, Inc.*, 2005.

[34] J.A. Russell, A. Weiss, and G.A. Mendelsohn. Affect-grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, 57:493–502, 1989.

[35] J. Scheirer, R. Fernandez, J. Klein, and R.W. Picard. Frustrating the user on purpose: A step toward building an affective computer. *Interacting With Computers*, 14:93–118, 2002.

[36] G. Turpin. Effects of stimulus intensity on automatic responding: The problem of differentiating orienting and defense reflexes. *Psychophysiology*, 23:1–14,, 1986.
[37] J. Wagner, J. Kim, and E. André. From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification. In *IEEE International Conference in Multimedia and Expo*, pages 940–943, 2005.
[38] R.D. Ward and P.H. Marsden. Physiological responses to different web page designs. *International Journal of Human-Computer Studies*, 59:199–212, 2003.
[39] G. N. Yannakakis and J. Hallam. Modeling and augmenting game entertainment through challenge and curiosity. *International Journal on Artificial Intelligence Tools*, 16:981–999, 2007.
[40] G. N. Yannakakis and J. Hallam. Entertainment modeling through physiology in physical play. *International Journal of Human-Computer Studies*, 66:741–755, 2008.
[41] G. N. Yannakakis and J. Hallam. Real-time game adaptation for optimizing player satisfaction. *IEEE Transactions on Computational Intelligence and AI in Games*, 1:121–133, 2009.
[42] L.A. Zadeh. Fuzzy sets. *Information Control*, 8:338–358, 1965.