

Characterising Players of a Cube Puzzle Game with a Two-level Bag of Words

Xavier Anadón
Pablo Sanahuja
Universitat Jaume I
Castellón, Spain

V. Javier Traver
Angeles Lopez
Jose Ribelles
vtraver@uji.es
Institute of New Imaging Technologies
Universitat Jaume I
Castellón, Spain

ABSTRACT

This work explores an unsupervised approach for modelling players of a 2D cube puzzle game with the ultimate goal of customising the game for particular players based solely on their interaction data. To that end, user interactions when solving puzzles are coded as images. Then, a feature embedding is learned for each puzzle with a convolutional network trained to regress the players' completion effort in terms of time and number of clicks. Next, the known bag-of-words technique is used at two levels. First, sets of puzzles are represented using the puzzle feature embeddings as the input space. Second, the resulting first-level histograms are used as input space for characterising players. As a result, new players can be characterised in terms of the resulting second-level histograms. Preliminary results indicate that the approach is effective for characterising players in terms of performance. It is also tentatively observed that other personal perceptions and preferences, beyond performance, are somehow implicitly captured from behavioural data.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning**; • **Information systems** → **Personalization**; • **Human-centered computing** → **User models**.

KEYWORDS

videogame, player characterisation, performance prediction, machine learning, clustering, bag of words

ACM Reference Format:

Xavier Anadón, Pablo Sanahuja, V. Javier Traver, Angeles Lopez, and Jose Ribelles. 2021. Characterising Players of a Cube Puzzle Game with a Two-level Bag of Words. In *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '21 Adjunct)*, June 21–25, 2021, Utrecht, Netherlands. ACM, New York, NY, USA, 7 pages. <https://doi.org/10.1145/3450614.3461690>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

UMAP '21 Adjunct, June 21–25, 2021, Utrecht, Netherlands

© 2021 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8367-7/21/06...\$15.00

<https://doi.org/10.1145/3450614.3461690>

1 INTRODUCTION

In a previous work [26], we explored how visual modification in a cube puzzle game could modulate the game challenge without any other modification of the gameplay whatsoever. Through a between-subject protocol, it was found not only that different visual modifications generally induced different challenges, but also, and more importantly, that even the more challenging modifications were perceived differently by different players. This is by no means a surprising result; after all, different people have different background and skills. But, motivated by the results and observations in this particular game, a natural question emerged: could players of this game be characterised from their interactive behaviour while playing so that the visual modifications of the game can be adaptive?

This work addresses the player characterisation part, not the game adaptation part. An overview of the proposed approach is given in Fig. 1. The problem is tackled through an unsupervised scheme to discover player profiles from behavioural data, instead of relying on ground-truth user classes that could be inferred from either measurements of self-reported personality traits or implicit association tests [28]. Also, instead of manually extracting features as in other works [12], we encode the players' interaction as images, and then learn deep features for a supervised regression task. Although the regression task is supervised, the rest of the approach is unsupervised in that these features are directly used for user modelling without additional ground-truth, predefined labels of puzzles or player profiles. After our background on the bag-of-words (BoW) representation applied to human action and gesture recognition [1, 32], and partly inspired by the "bag of behaviours" [7] in a different user modelling problem and with a different purpose, we propose to use a two-level BoW: the first one will characterise sets of individual puzzles; and the second one will characterise players.

The main contribution of this work is an approach for characterising players of a particular videogame, which combines known computer vision and machine learning techniques. Although the methodology and tests are focused on a case-study game as part of an ongoing project, some of the ideas and concepts underlying our proposal are likely to be applicable to some other games as well. In particular, games with significant visual components and mouse- or touch-based interactions might find useful the possibilities of encoding interactions with images, feature learning, or some approach similar to the bag-of-words representation.

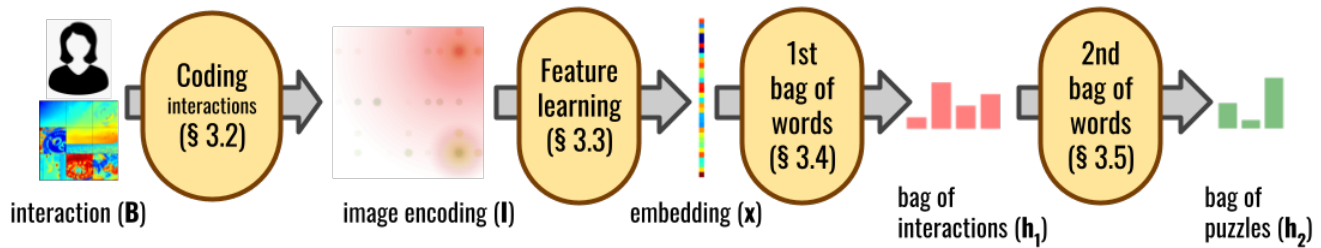


Figure 1: Schematic overview of the proposed methodology.

2 RELATED WORK

Player modelling. Modelling players of videogames is useful for marketing, procedural content generation, and game design. Dynamic difficulty adjustment (DDA) [39] is particularly relevant to adapt the challenge to the players’ skills or preferences, which may bring an improved game experience [35]. Personality and game experience are related [5]; player types can be found from personality traits, and the latter can be inferred through behavioural data [9], which in turn can help personalise games [16, 31]. User modelling from either theoretical frameworks [30], or human domain experts [20], are useful, but have limited flexibility. Naturally, data-driven approaches pose an interesting alternative with advantages such as generalisation to other games [14]. Recently, player behaviours are modelled with neural networks, which can be applied to generating the behaviour of the opponent player so as to modulate the difficulty [23, 24].

Predictive systems. Player experience (challenge, frustration, and fun) can be modelled through controllable features of level design [21, 22]. Instead of having users report their personality or emotions explicitly during the game play, it is interesting to do this unsupervisedly and implicitly [4], or through gaze, and physiological data [19]. Sequential models of in-game player behaviours are an alternative to aggregated players’ actions, for predicting personality and expertise [6], assistance in serious games [36], churn prediction [17, 38], or player categorisation from past and predicted behaviours [8]. Deep and reinforcement learning may help predict completion rate [18], or excessive gaming [34].

Challenging scenarios. In some cases, data from players is simply unavailable. For modelling players who leave the game early, data from both, other players of the same game, and other games played by the target player, are explored via transfer learning [33]. Computational models of motivation, and artificial game-playing agents have been proposed [13] to predict player experience without any actual player. Although in-game data is crucially important to model players [27], this data is not available during game development, and AI-based players can be used instead [11]. Our work uses data from actual players, but future work might consider how to include data from computational players.

3 METHODOLOGY

After describing the puzzle game and the user study from which behavioural data is obtained (Sec. 3.1), we introduce how the interaction of one player with one puzzle can be encoded as an image

(Sec. 3.2). These images are then used to learn to predict the player performance via supervised regression (Sec. 3.3). As a byproduct of this learning, interactions can now be represented with a compact feature vector which is used for the two-level bag-of-words approach (Sec. 3.4, Sec. 3.5).

3.1 Interaction data from a cube puzzle game

In this work, we use the data gathered from a case study conducted in a previous work [26], which explored how visual modifications in a particular game could modulate the game challenge. The case study consisted of a web-based cube puzzle game (Fig. 2) with two different versions. The experiment was conducted online, with a between-subject protocol, where each participant played only one version, randomly assigned. We collected behavioural data for each player, as well as their responses to a final questionnaire about their opinion, preferences and perceived effort (Table 1). Now, in this work, we aim at characterising players of this game from the collected data.

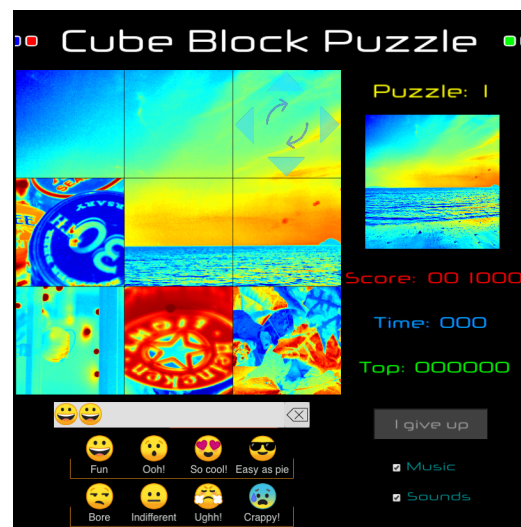


Figure 2: One puzzle in the VC version for the beach target image, colour modification and distracting images unrelated to the target image. Five out of the nine cubes have already the correct side on.

In this cube puzzle game, participants had to solve six cube puzzles, sequentially presented to them. Each puzzle consisted of

Table 1: After-game questionnaire

1	In general, I found easy to complete the game	8	I found the overall experience to be exciting/frustrating/NA
2	I found the game entertaining	9	Which puzzle did you like the most?
3	I think I was quick solving the puzzles	10	Which puzzle did you like the least?
4	I would play this game again	11	I liked this game more than Mahjong
5	I found the overall experience to be entertaining/boring/NA	12	I liked this game more than Solitaire
6	I found the overall experience to be simple/complex/NA	13	I liked this game more than classical puzzle
7	I found the overall experience to be surprising/dull/NA	14	I liked this game more than Sliding puzzle

nine cubes, with only one cube side visible at a time (Fig. 2). Six images are involved in each puzzle, one per cube side. One of these images is the target image, and is displayed as a reference to the right of the cubes (right-hand side in Fig. 2). The participants had to form the target image through mouse clicks in delimited areas of each cube. These areas are hinted as arrows when hovering, as can be appreciated at the upper-right square of the puzzle in Fig. 2. These clicks are mapped to associated cube rotations around three orthogonal axes; namely, left-right arrows for pan, up-down arrows for tilt, and inner arrows for roll.

Players had also the choice to provide instant emotional feedback regarding each puzzle, in the form of emojis, which are available in the lower part of the the window (Fig. 2), but since this information was used very little, it is not considered in this work.

We used three different target images (eye, beach, smoke) along with five other images to design the six puzzles of the game. Two versions of the game were produced: the standard (ST) one, consisting of the six puzzles with the original images, and the visual computing (VC) one, formed by the same puzzles with all their images altered with a single visual concept. We used three visual concepts: edges (spatial gradients of the gray-level images), colour transformation (by applying a colour map), and dynamic transformation (by clockwise rotation of images). Each visual concept was applied to the six images (one per cube side), each in two out of the six puzzles (i.e. 3 concepts \times 2 = 6 puzzles). For further details on the game and the case study, the reader is referred to [26].

In this work, we use performance data from the 126 (ST and VC) participants who completed the game and the questionnaire. These data from each player at each puzzle are referred to as an *interaction* (represented as **B** in Fig. 1). Therefore, each interaction is a stream of time-stamped information about the mouse position and clicks which thus represent the trajectory as well as the sequence of cube rotations performed by a single player to solve one particular puzzle.

3.2 Coding interactions with images

The information of the interaction of a player solving one puzzle (**B**) is sequential in nature due to the temporal order of mouse movements and clicks. However, this information can be somehow coded as a single image as well. This idea is similar to how other sequential information has been represented in other problems such as coding audio information [2] or mouse movements in web search tasks [3].

Two different image encodings were initially considered: trajectory-based and click-based. In both cases the time information is colour-coded using a map colour from green to red, scaling time relatively

to the range $[0, 1]$, i.e. not using absolute time values. In the trajectory case, intermediate mouse positions are joined by line segments and time is additionally coded as the width of these line segments. In the case of clicks, each click is represented by a circle, whose position is the center of the corresponding arrow button. Since more than one click on the same arrow button is possible, the size of each circle is made proportional to the number of clicks. By using transparency, overlapping circles are still partially visible. Examples of these images are given in Fig. 3 for illustration purposes. In this work, we focus on the click-based image representation since it appeared to better predict performance in some early experiments. This procedure produces the image (**I** in Fig. 1) encoding an interaction.

3.3 Learning to predict performance

The $3 \times 512 \times 512$ colour images encoding the interactions were used to train a convolutional neural network (CNN) for regressing the performance (number of clicks and completion time). This prediction is not an end in itself because, after all, if we can construct these images, we can also know the values for the time and number of clicks. However, using a CNN for this task has two valuable purposes. First, we can find out how successful a CNN can be for predicting information from this type of images. Second, we can use the activation of one hidden layer to compactly represent the input image for subsequent tasks. In a way, this performance prediction resembles a form of self-supervised learning task [10].

For that regression purpose, a ResNet34 was used as the backbone CNN as a reasonable tradeoff between model complexity and prediction performance. Since ResNet was trained on a 1000-class classification problem, the last fully-connected (FC) classification layer was removed and replaced by two blocks each consisting of batch normalisation, one dropout layer (with drop rates $p = 0.2$ and $p = 0.5$, in each block, respectively), and one FC layer. The FC layer in the first block has 128 units (so that dimensionality is progressively reduced, as customary), and was followed by a ReLU activation. The output of the network at this point is used for the feature embedding $\mathbf{x} \in \mathcal{R}^{128}$. The FC layer of the second block has two units, corresponding to the predicted time and click values, respectively, and no activation function was used. The loss function was the mean squared error. A batch size of 32 instances were used.

We use the weights corresponding to training with ImageNet, so the corresponding mean and standard deviations of the colour channels for the training images was used to normalise our input images. For training, we first freeze these weights and train the new layers for 40 epochs with the Adam optimiser, and learning rate $5 \cdot 10^{-4}$. Then, the ImageNet-pretrained weights are unfrozen

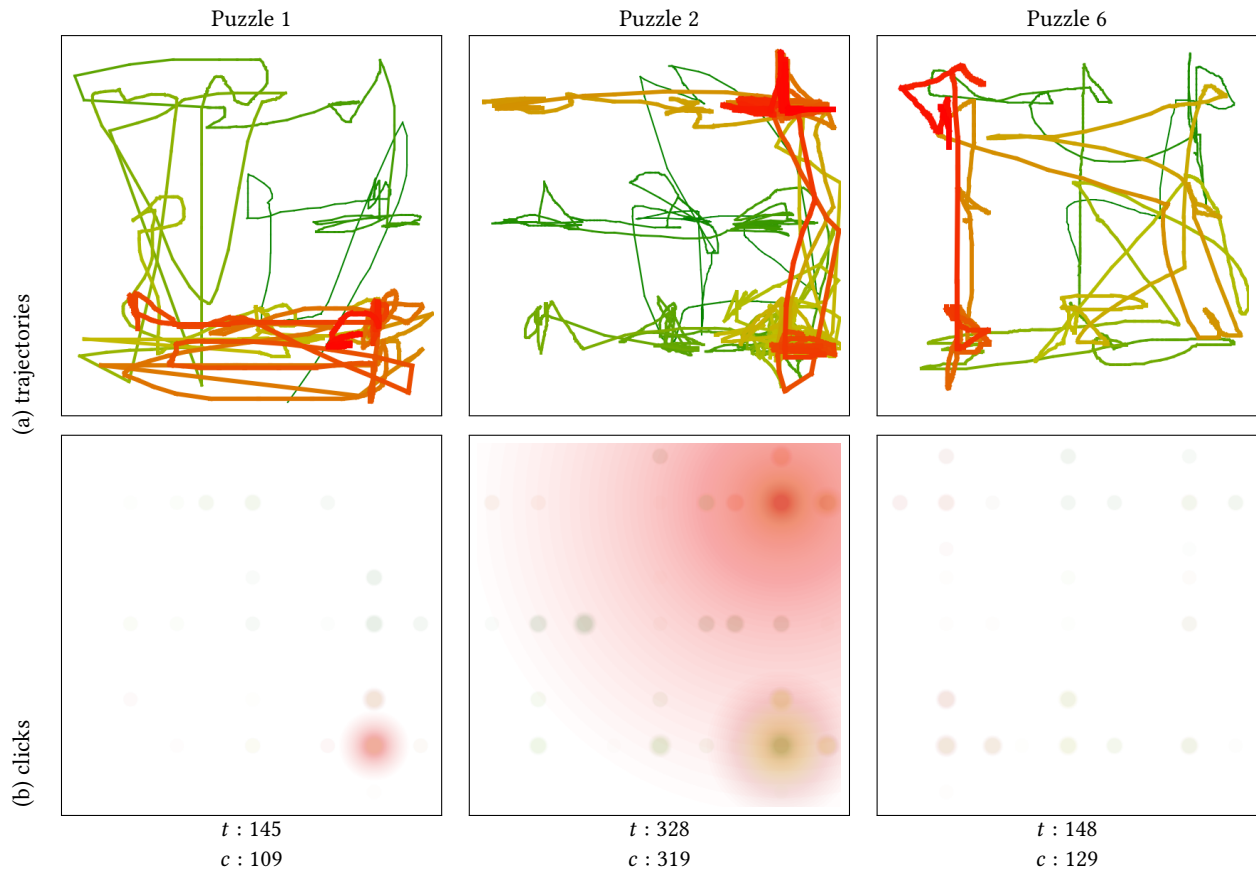


Figure 3: Examples of images encoding behavioural data per puzzle: (a) trajectory-based and (b) click-based, for three different users and puzzles. In each case, the elapsed time (t) in seconds, and the number of clicks (c) are given below, as a reference.

and the full network is trained for 10 more epochs with a lower learning rate of 10^{-5} , so as to get some further adaptation to this specific task.

3.4 Characterising puzzles with bags of interactions

To characterise behaviours of players with individual puzzles, sets of puzzles and, ultimately, players, we use the well-known concept of the bag of words (BoW) [29]. In essence, the BoW consists of vector quantization (i.e. clustering a set of feature vectors) to build a vocabulary (the dictionary of words), and using a histogram (i.e. the bag of words) as a pooling mechanism to summarise a given document (i.e. a set of words). The vocabulary is typically represented by the centroids of the C clusters, with C being the chosen size of the vocabulary. Here, k -means was used for clustering, and the number of clusters (vocabulary size) was manually selected using as a guiding criteria the well-known elbow method [37] from the set of tested $k \in \{1, \dots, 9\}$.

In our case, the vocabulary is built from the training set of interactions represented by the corresponding embedding vectors. Now, a single new interaction \mathbf{x} can be represented in terms of this vocabulary as $\mathbf{h}(\mathbf{x}) \in \mathbb{R}^C$ by using a soft cluster assignment.

Formally, let $\{d_i\}_{i=1}^C$ the distances of \mathbf{x} to each of the C clusters; the i -th bin of \mathbf{h} is computed by the following softmax function:

$$\mathbf{h}_i(\mathbf{x}) = \frac{e^{-d_i}}{\sum_{j=1}^C e^{-d_j}}, \quad (1)$$

which is similar to the idea of fuzzy c -means [15]. Soft assignments are generally beneficial [25]. Note that this soft-assignment of a single vector to the set of clusters can be seen as a histogram itself. Now, the BoW representation for a set $S = \{\mathbf{x}_j\}_{j=1}^N$ of embeddings, corresponding to a set of N puzzles (interactions), can be simply computed as the sum

$$\mathbf{h}(S) = \sum_{j=1}^N \mathbf{h}(\mathbf{x}_j), \quad (2)$$

and then normalised with L_1 . Both the single-puzzle histogram (Eq. 1) and the multi-puzzle histogram (Eq. 2) correspond to the first-level bag of words (\mathbf{h}_1 in Fig. 1). The left part of Fig. 4 summarises the procedure.

3.5 Characterising players with bags of puzzles

The BoW discussed above (Sec. 3.4) is performed from embeddings corresponding to *individual* puzzles. To model players, however, a

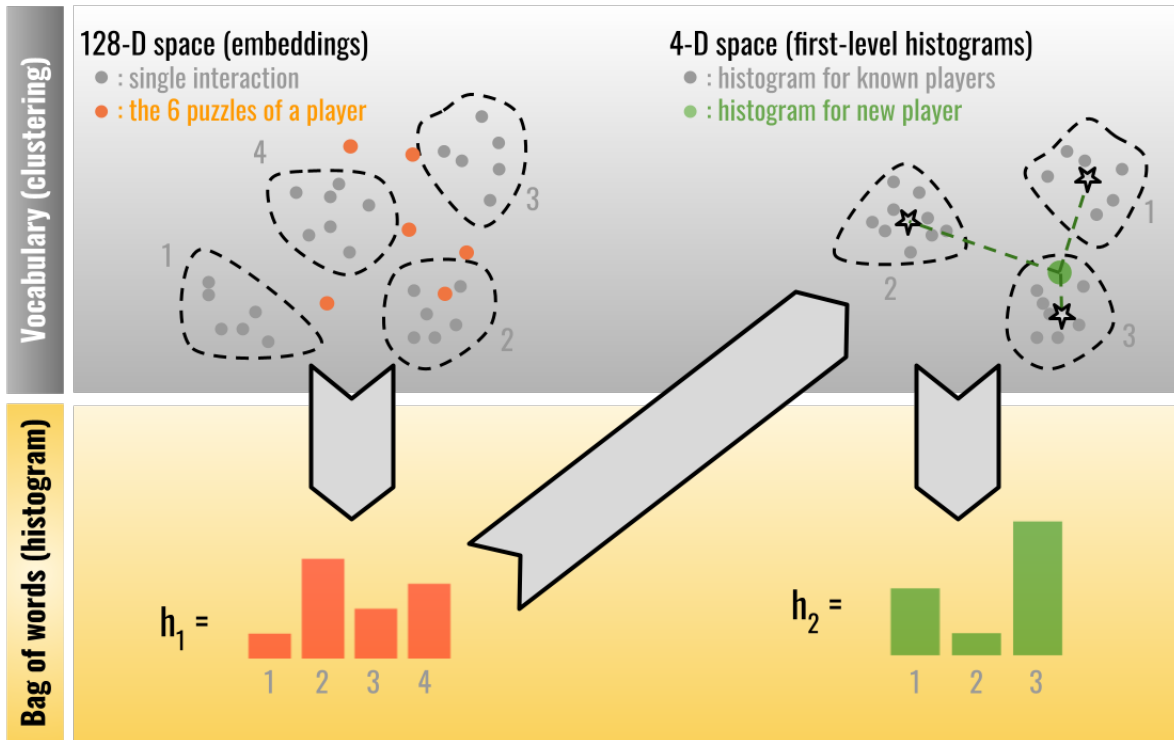


Figure 4: The two-level bag-of-words illustrated. the two levels are similar and involve two steps each: building a vocabulary via a clustering procedure (upper part), and defining the bag-of-words through histogramming (lower part). The difference between the two levels is what are the input data points and what the resulting histograms represent. Left: The first level (Sec. 3.4) has as input the feature vectors learned in the regression CNN (Sec. 3.3), and the resulting histograms represent sets of interactions. Right: The second level (Sec. 3.5) uses the histograms produced from the first-level BoW, each corresponding to the set of interactions from a single player, and produces histograms, each of which represent how a given (unknown or new) player is similar to each of the identified groups (clusters) of players. Both BoW levels use a soft assignment (Eq. 1) based on the distances (dashed line segments) of a data point to the cluster centres (stars). The number clusters in both cases were identified with the elbow method.

conceptually higher stage is required. Since a single player i with a set of puzzles S_i for all its interactions can be represented by $\mathbf{h}(S_i)$, it can be argued that profiles of players can now be discovered by operating in the space of these histograms of *joint* puzzles. Therefore, we propose to perform another BoW, this time using the histograms of the first-level BoW (\mathbf{h}_1) as an input. In turn, the resulting vocabulary will serve to characterise players in terms of the second BoW (\mathbf{h}_2 in Fig. 1), which can be useful to assign a new player to one profile, or to a mixture of profiles, given a complete or partial set of their interactions. The procedure is illustrated in the right part of Fig. 4. The number of clusters was again chosen by the elbow method from the same set of tested k values as in the first level.

4 RESULTS

From the second-level BoW, three player groups are automatically identified. As shown in Fig. 5, these groups roughly correspond with their actual effort. In particular, Group 1 correspond to the players with larger times and more clicks. Interestingly, it can

tentatively be argued that Groups 2 and 3 have similar times, but differ in the number of clicks. Although the group separation is not perfect, possibly due to limited data or learning, this result tentatively indicates that (1) the image-based representations of the interactions encode rich information regarding player’s behaviour; (2) the learned (128-dimensional) feature embeddings found by the CNN captures compactly the player performance; and (3) the BoW representation is able to distinguish patterns of players from the feature embeddings.

Although feature embeddings were learned for predicting times and clicks, they might also be latently capturing other player’s characteristics since the input images are representing specific interactive behaviours. To explore this, we looked into the players’ responses to the questionnaire. As seen in Fig. 5, players in Group 1 took the longest to complete the game, and even though they were aware of it (Fig. 6-Q₃), they would play the game again (Fig. 6-Q₄). Although a finer-grained analysis would be required, it can roughly be stated that players within this profile could be offered to play similar puzzles in the future since their skill and puzzle challenge seem to be aligned. On the other hand, even though players in

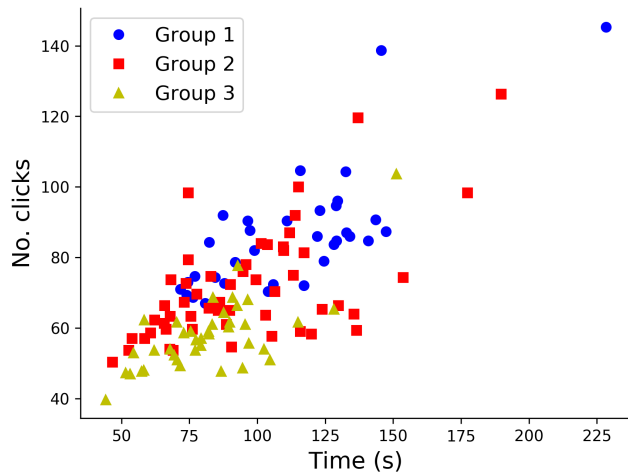


Figure 5: Average completion times and number of clicks per puzzle for the three player’s groups automatically identified in the second-level BoW. Each point correspond to a different player. Although the proposed approach allows for a soft assignment whereby a single player has a degree of membership to each player group, here a hard assignment to the closest cluster has been used.

Table 2: Classification of players into groups from their questionnaire responses. Number of players (column-wise % in parentheses).

pred.\true	1	2	3
1	15 (44.1)	12 (22.6)	13 (33.3)
2	8 (23.5)	23 (54.8)	11 (28.2)
3	11 (32.3)	18 (34.0)	15 (38.5)
	34 (100)	53 (100)	39 (100)

Group 3 took the least to complete the game (Fig. 5), and this aligns with their subjective perception (Fig. 6-Q₁, Fig. 6-Q₃), they were the least to think they would play again (Fig. 6-Q₄). This suggests that these players might have found the game boring, arguably because their skills are higher than the game challenge. Tentatively, harder puzzles could probably be suitable for them.

Finally, we explored how much the responses to the questionnaire (Q₁–Q₁₄, Table 1) relate to players profiles.

Results with the nearest-neighbour classifier and leaving-one-player-out (Table 2) indicate that, despite the unsurprising low overall accuracy (42.1%), it happens that within each true group, the predicted group with most players correspond to the correct one, which suggests that simple behavioural data might subtly capture some subjective traits beyond performance.

5 DISCUSSION

Although the click-based image encoding was used in this work, the trajectory-based representation could also be explored, since it may model intrinsic players’ skills, e.g. in terms of puzzle solving strategies. Since the current data size is somehow limited, collecting

data for more players would be helpful. At this stage, we preferred to keep the approach mostly unsupervised, but embeddings could also be learned for supervisedly learning preferences and perceptions, which could provide complementary predictive cues. The image-based representation is very suitable for CNNs, but it does not easily lend itself to on-line predictions for earlier user characterisation (e.g. before completing a puzzle); exploring long short-term memories (LSTMs) for this problem, either from image encoding or raw mouse data, seems a natural next step.

6 CONCLUSION

A framework for characterising players of a cube puzzle game has been proposed. It relies on representing the behavioural interaction as images, and then learning a feature embedding. These compact learned feature vectors are then used for representing individual and sets of puzzles as histograms following a two-level bag-of-words approach. Preliminary results suggest this pipeline is appropriate for comparing puzzles, and players in terms of the puzzles they played, which offers a simple and reasonably effective strategy for player characterisation in terms of performance and beyond.

REFERENCES

- [1] Pau Agustí, V. Javier Traver, and Filiberto Pla. 2014. Bag-of-words with aggregated temporal pair-wise word co-occurrence for human action recognition. *Pattern Recognition Letters* 49 (2014), 224–230.
- [2] Relja Arandjelovic and Andrew Zisserman. 2018. Objects that Sound. In *Proceedings of the European Conference on Computer Vision (ECCV)*.
- [3] Ioannis Arapakis and Luis A. Leiva. 2020. Learning Efficient Representations of Mouse Movements to Predict User Attention. In *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*.
- [4] Preeti Bhargava, Oliver Brdiczka, and Michael Roberts. 2015. Unsupervised Modeling of Users’ Interests from Their Facebook Profiles and Activities. In *Proc. of the 20th Intl. Conf. on Intelligent User Interfaces (IUI ’15)*. 191–201.
- [5] Max V. Birk, Dereck Toker, Regan L. Mandryk, and Cristina Conati. 2015. Modeling Motivation in a Social Network Game Using Player-Centric Traits and Personality Traits. In *User Modeling, Adaptation and Personalization*, Francesco Ricci, Kalina Bontcheva, Owen Conlan, and Séamus Lawless (Eds.). 18–30.
- [6] Sara Bunian, Alessandro Canossa, Randy C. Colvin, and Magy Seif El-Nasr. 2017. Modeling Individual Differences in Game Behavior Using HMM. In *Proceedings of the Thirteenth Conference on Artificial Intelligence (AAAI) and Interactive Digital Entertainment (AIIDE)*, Brian Magerko and Jonathan P. Rowe (Eds.). AAAI Press, 158–164.
- [7] Yuya Chiba, Masashi Ito, Takashi Nose, and Akinori Ito. 2014. User Modeling by Using Bag-of-Behaviors for Building a Dialog System Sensitive to the Interlocutor’s Internal State. In *Proceedings of the Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*. 74–78.
- [8] A. F. del Río, P. Pei Chen, and Á. Periañez. 2019. Profiling Players with Engagement Predictions. In *2019 IEEE Conference on Games (CoG)*. 1–4.
- [9] Mouna Denden, Ahmed Tlili, Fathi Essalmi, and Mohamed Jemni. 2018. Implicit modeling of learners’ personalities in a game-based learning environment using their gaming behaviors. *Smart Learning Environments* 5, 1 (2018).
- [10] C. Doersch, A. Gupta, and A. A. Efros. 2015. Unsupervised Visual Representation Learning by Context Prediction. In *IEEE Intl. Conf. on Computer Vision (ICCV)*. 1422–1430.
- [11] A. Feng and A. S. Gordon. 2020. Recognizing Multiplayer Behaviors Using Synthetic Training Data. In *2020 IEEE Conference on Games (CoG)*. 463–470.
- [12] J. Gow, R. Baumgarten, P. Cairns, S. Colton, and P. Miller. 2012. Unsupervised Modeling of Player Style With LDA. *IEEE Transactions on Computational Intelligence and AI in Games* 4, 3 (2012), 152–166.
- [13] Christian Guckelsberger, Christoph Salge, Jeremy Gow, and Paul Cairns. 2017. Predicting Player Experience without the Player: An Exploratory Study. In *Proc. of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY ’17)*. 305–315.
- [14] Daniaal Hooshyar, Moslem Yousefi, and Heuseok Lim. 2018. Data-Driven Approaches to Game Player Modeling: A Systematic Literature Review. *Comput. Surveys* 50, 6 (Jan. 2018).
- [15] Anil K. Jain. 2010. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters* 31, 8 (2010), 651–666.

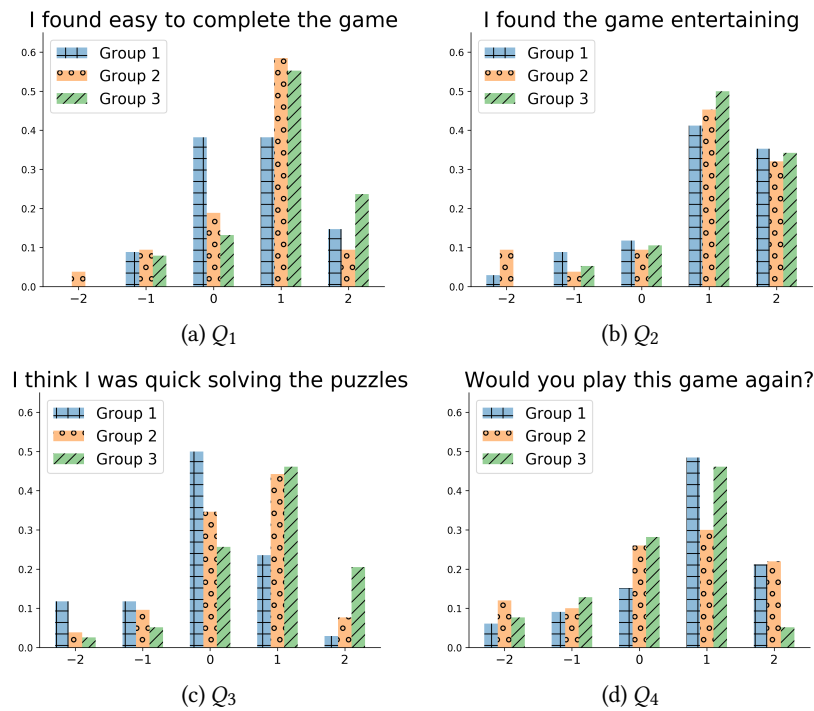


Figure 6: Relative histograms of 5-point Likert responses (from -2 =“strongly disagree” to 2 =“strongly agree”) to some items Q_i of the questionnaire.

[16] Stephen Karpinskyj, Fabio Zambetta, and Lawrence Cavedon. 2014. Video game personalisation techniques: A comprehensive survey. *Entertainment Computing* 5, 4 (2014), 211–218.

[17] J. T. Kristensen and P. Burelli. 2019. Combining Sequential and Aggregated Data for Churn Prediction in Casual Freemium Games. In *2019 IEEE Conference on Games (CoG)*. 1–8.

[18] J. T. Kristensen, A. Valdivia, and P. Burelli. 2020. Estimating Player Completion Rate in Mobile Puzzle Games Using Reinforcement Learning. In *2020 IEEE Conference on Games (CoG)*. 636–639.

[19] S. Lee-Cultura, K. Sharma, S. Papavaslopoulou, and M. Giannakos. 2020. Motion-Based Educational Games: Using Multi-Modal Data to Predict Player’s Performance. In *2020 IEEE Conference on Games (CoG)*. 17–24.

[20] R. Mathonat, J. F. Boulicaut, and M. Kaytoue. 2020. A Behavioral Pattern Mining Approach to Model Player Skills in Rocket League. In *2020 IEEE Conference on Games (CoG)*. 267–274.

[21] Chris Pedersen, Julian Togelius, and Georgios N. Yannakakis. 2009. Modeling Player Experience in Super Mario Bros. In *Proceedings of the 5th International Conference on Computational Intelligence and Games*.

[22] C. Pedersen, J. Togelius, and G. N. Yannakakis. 2010. Modeling Player Experience for Content Creation. *IEEE Transactions on Computational Intelligence and AI in Games* 2, 1 (2010), 54–67.

[23] J. Pfau, A. Liapis, G. Volkmar, G. N. Yannakakis, and R. Malaka. 2020. Dungeons & Replicants: Automated Game Balancing via Deep Player Behavior Modeling. In *2020 IEEE Conference on Games (CoG)*. 431–438.

[24] Johannes Pfau, Jan David Smeddinck, and Rainer Malaka. 2019. Deep Player Behavior Models: Evaluating a Novel Take on Dynamic Difficulty Adjustment. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland UK) (CHI EA ’19)*.

[25] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. 2008. Lost in quantization: Improving particular object retrieval in large scale image databases. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*.

[26] Jose Ribelles, Angeles Lopez, and V. Javier Traver. [n.d.]. Modulating the gameplay challenge through simple visual computing elements: A cube puzzle case study. (*Submitted*) (n.d.).

[27] Jacques D. Rodrigues, Luiz A. L. Brancher. 2018. Improving Players’ Profiles Clustering from Game Data Through Feature Extraction. In *Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*.

[28] Carlos Pereira Santos, Kevin Hutchinson, Vassilis-Javed Khan, and Panos Markopoulos. 2019. Profiling Personality Traits with Games. *ACM Trans. Interact. Intell. Syst.* 9, 2–3 (March 2019).

[29] J. Sivic, B. C. Russell, A. A. Efros, A. Zisserman, and W. T. Freeman. 2005. Discovering objects and their location in images. In *IEEE Intl. Conf. on Computer Vision (ICCV)*, Vol. 1. 370–377.

[30] Sam Snodgrass, Omid Mohaddesi, Jack Hart, Guillermo Romera Rodriguez, Christoffer Holmgård, and Casper Hartevelde. 2019. Like PEAS in PoDS: The Player, Environment, Agents, System Framework for the Personalization of Digital Systems. In *Proceedings of the International Conference on the Foundations of Digital Games*.

[31] A. Tlili, M. Denden, F. Essalmi, M. Jemni, Kinshuk, N. Chen, and R. Huang. 2019. Does Providing a Personalized Educational Game Based on Personality Matter? A Case Study. *IEEE Access* 7 (2019), 119566–119575.

[32] V. Javier Traver, Pedro Latorre Carmona, Eva Salvador-Balaguer, Filiberto Pla, and Bahram Javidi. 2017. Three-Dimensional Integral Imaging for Gesture Recognition Under Occlusions. *IEEE Signal Processing Letters* 24, 2 (2017), 171–175.

[33] Andrej Vitek. 2020. Cross-Game Modeling of Player’s Behaviour in Free-To-Play Games. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP ’20)*. 384–387.

[34] Q. Wu and J. Carrette. 2020. Can Deep Learning Predict Problematic Gaming?. In *2020 IEEE Conference on Games (CoG)*. 662–665.

[35] Su Xue, Meng Wu, John Kolen, Navid Aghdaie, and Kazi A. Zaman. 2017. Dynamic Difficulty Adjustment for Maximized Engagement in Digital Games. In *Proc. of the Intl. Conf. on World Wide Web Companion (WWW ’17 Companion)*. 465–471.

[36] Vivek Yadav, Alexander Streicher, and Ajinkya Prabhune. 2020. User Assistance for Serious Games Using Hidden Markov Model. In *Addressing Global Challenges and Quality Education - European Conference on Technology Enhanced Learning, EC-TEL*. 380–385.

[37] Chunhui Yuan and Haitao Yang. 2019. Research on K-Value Selection Method of K-Means Clustering Algorithm. *J* 2, 2 (2019), 226–235.

[38] S. Zhao, R. Wu, J. Tao, M. Qu, H. Li, and C. Fan. 2020. Multi-source Data Multi-task Learning for Profiling Players in Online Games. In *2020 IEEE Conference on Games (CoG)*. 104–111.

[39] Mohammad Zohaib and Hideyuki Nakanishi. 2018. Dynamic Difficulty Adjustment (DDA) in Computer Games: A Review. *Advances in Human-Computer Interaction* (Jan. 2018).