Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 15 (2012) 140 – 147

Procedia
Computer Science

[*]Virtual Worlds for Serious Applications (VS-GAMES'12)

Towards Detecting Clusters of Players using Visual and Gameplay Behavioral Cues

Stylianos Asteriadis[a,*], Kostas Karpouzis[a], Noor Shaker[b], Georgios N. Yannakakis[b]

[a]Image, Video and Multimedia Systems Lab, National Technical University of Athens, 9 Ir. Polytechniou str, GR-157 80, Athens, Greece
[b]IT University of Copenhagen, Rued Langaards Vej 7, 2300 Copenhagen, Denmark

Abstract

The issue of discriminating among players' styles and associating them with player profile characteristics, demographics and specific interests and needs is of vital importance for creating content, fine tuned and optimized in such a way that user engagement and interest are maximized. This paper attempts to address the issue of clustering players' behavior using visual features and player performance, as input parameters. Following an unsupervised scheme, in this work, we utilize data from Super Mario game recordings and explore the possibility of retrieving classes of player types along with existing correlations with certain global characteristics.

Keywords: Player Modelling; Expressivity during gameplay; Motion tracking; Game Context;

1. Introduction

Identifying different player styles, preferences and characteristics in games is receiving a lot of attention in the last years. Delivering game content matching the particular player profile, preferences and demographics is a key factor for the game industry, in order to develop personalized and highly adaptable content to the customers. A typical strategy usually followed by developers is that of interviews or questionnaires designed to unlock cues related to player's personality [1], a technique that has the risk of subjectivity and leads to time-consuming procedures. In this paper, different player clusters (styles) are distinguished based on play and visual behavior. The relation of demographics and game experience with overall visual activation, reactions

[*] Corresponding author. Tel.: +30-210-7724052; fax:30-210-772-2492.
E-mail address: stiast@image.ntua.gr.

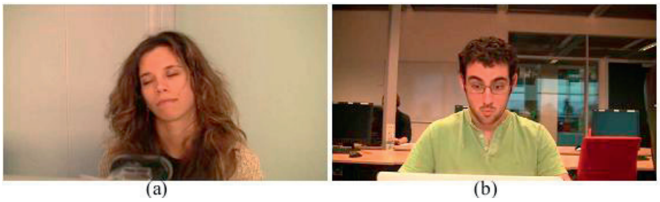
1877-0509 © 2012 The Authors. Published by Elsevier B.V. Selection and/or peer-review under responsibility of the scientific programme committee of VS-Games 2012
doi:10.1016/j.procs.2012.10.065

during critical events and player actions is being explored through unsupervised algorithms used here to discover dynamics and hidden structures related to player types.

In recent bibliography, various works have been published, dealing with the issue of exploring different player styles, preferences and classes. From Massively Multiplayer Online Games (MMOGs), like World of Warcraft (WoW) to Single Player games, like Super Mario, there is a large span of properties that can be taken into account for modeling player's behavior, style, preferences, social attitude and demographics. Authors in [2], for instance, study the impact of the social character of WoW to the play behavior, and examine those features that players concentrate on, suggesting re-designing certain game aspects, in accordance to their findings. In [3], Hidden Markov Models are employed to model time dynamics of player actions in a version of Zereal [4], in order to achieve robust classification of player types in terms of game preferences. Using a large dataset of play habits collected online, Yee et al. in [5] explore those demographic characteristics that are linked to certain behaviors. Thus, by following a multiple regression methodology, they link certain features (age, gender, experience) to preferences and gameplay activities. Focusing on the Five Factor Model of personality (FFM), authors in [6] analyze the behavior of 44 participants at a role-play game and explore the impact of personality traits to game behavior. They report correlations between game behavior and all five personality dimensions of the FFM (Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism). Working on a first-person social hunting game, authors in [7] introduce a meta-clustering scheme, separating classification into three levels: Action/Skill based, Preference based and Socially based. The above are structured in a hierarchical manner, in order to be combined and associated with different player styles. Analytical relations among the resulted structures allowed the authors to propose design mechanisms aiming at boosting social interaction among players.

Using player demographics, levels of arousal and visual feedback, fused together, can also act as a reliable indicate [8-9] with works mostly focusing on estimating player state, preferences or emotions. In [8], for example, the authors use electrodermal activity along with heart rate to establish correlations with self-reported data. Asteriadis et al in [9] combine visual behavior with player profile characteristics in order to reveal the potentiality of estimating levels of reported Challenge, Frustration and Engagement. In this work, we propose a framework for identifying player types based on their visual behavior and total play time. By using visual feedback, this work goes one step beyond previous research where, mostly, game actions and features were utilized for retrieving player classes. We employ a clustering scheme and choose the optimal number of clusters based on certain criteria. By unfolding such categorizations, the aim is to examine which visual feedback and player performance cues can be related to certain player profile features. Exploring links between player profile and reactions/performance will allow to interpret certain expected or unexpected feedback from the player and, subsequently, manage game parameters accordingly. In this way, personalized or profile-dependent game adaptations can allow interest and engagement maximization, which is the final scope of any game environment. Authors in [9] have shown that it is possible to achieve high degrees of accuracy ($69-74\%$) at estimating engagement, challenge and frustration levels, as reported by the players themselves, by making use of head movements during certain events, as well as player characteristics (age, experience). In that work, however, different considerations of expressivity, according to different player features were not considered, proposing a generic model. The work presented in this paper is towards the direction of personalization and user modeling, however, building on different player types and explanations of certain behaviors.

Fig. 1. Snapshot from Infinite Mario Bros, showing Mario standing on horizontally placed boxes surrounded by different types of enemies



(a) (b)

Fig. 2. Facial Reactions of players at moments of losing

2. Data Acquisition

2.1. Game Environment

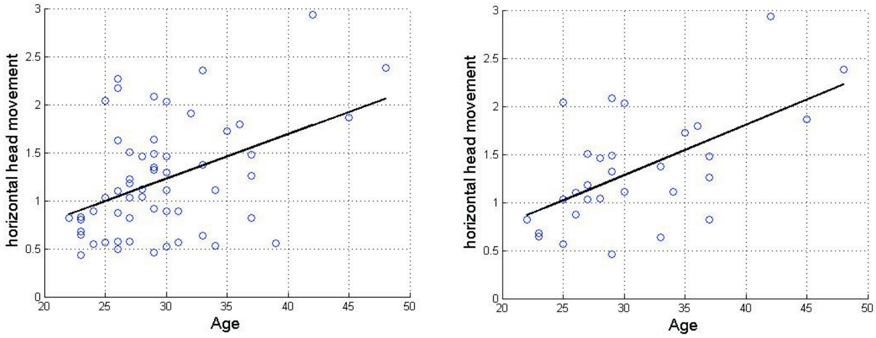
For extracting correlations between visual features (facial gestures) and interaction content, we used a modified version of Super Mario, which provides time-stamped output of game-related events (player jumping or changing direction, loss, level end, power up use, etc.) (Fig. 1). The description of the protocol followed for collecting data can be found in [10], while, in this study, the dataset was extended reaching a total of 58 participants (30 female, 28 male), corresponding to about 4.7 hours of game play material. People's upper body was recorded while they were playing, and certain facial gestures were measured during certain events (here, whenever Super Mario was killed, Fig. 2), as well as, throughout game play of a whole stage per player.

2.2. Facial Features Extraction

Head movement expressivity was captured and analyzed during game play. Initially, the face is detected based on a fusion of techniques: Viola-Jones [11] followed by an ellipse fitting algorithm [12] are used for a precise estimate of face location. A skin detection step is also used here, to get the precise contour of the face [13] and the upper, lower, left and rightmost points are extracted. At each frame, eye centers and eyes midpoints are detected using the methodology described in [12]. Subsequently, the euclidean distance of the eyes midpoint with regards to its initial position at the beginning of the game session is extracted and its first derivative is considered as the head movement. The visual input considered in this paper is the average horizontal component of head movement per player (H_{av}), as well as the average horizontal component of head movement she/he poses at instances of losing (H_l). The choice of horizontal movements, over average or vertical movements was preferred, due to players' tendency to pose horizontal movements more than towards other directions.

Initial analysis using regression models correlating the above features with critical events was performed, but, also, it was observed that performance is linked to expressiveness. More in particular, horizontal movements of head appear to have strong correlations with age at instances when Super Mario is killed (r=0.43, R^2=0.19, F=12.7, p=0.0007), as shown qualitatively in Fig. 3(a). Most of the expressivity, however, appears to be coming from women. Furthermore, this was more evident as age would increase. In particular, Fig. 3(b) shows the dependency of head horizontal movements on age, for women (r=0.58, R^2=0.34, F=14.4,

p=0.0007), during moments of critical events. Similar, significant correlations between movements during general gameplay with age for women were found (r=0.46, R^2=0.21, F=7.6, p=0.01).



(a) (b)

Fig. 3. (a) Horizontal head movements for all players and (b) women only, in relation to age, during instances when player loses (Super Mario is killed).

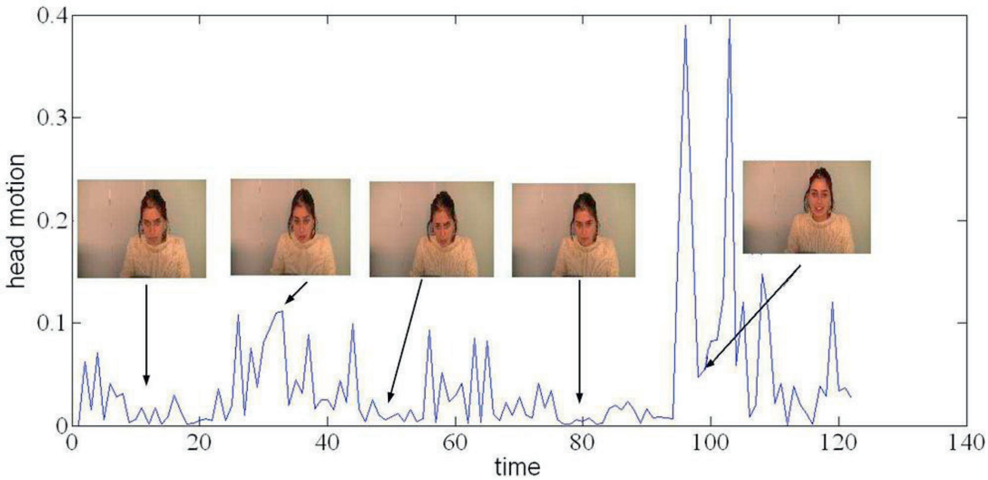


Fig. 4. Player visual behavior during gameplay. In this session, Super Mario was killed in seconds ≃ 32 and ≃ 100

3. Clustering approach

For all 58 players, the features used were average head movement, head movement when they lose and average time played per player. This last feature was employed as a way to express performance since the better a player is, the more time she/he will play before losing. The clustering approach followed was the k-medoids algorithm, thanks to its higher robustness to noise and outliers, in comparison to the classical k-means. The technique followed for the realization of k-medoids was the Partitioning Around Medoids (PAM) algorithm [14] which functions as follows: A set of randomly selected data points are assigned the role of medoids (centrally positioned points in clusters) and the rest of the data are clustered around the medoids they are closer to (here, we have used the common Euclidean distance). For defining the real medoids, roles between data and medoids are continuously swapped and the total cost of each configuration is calculated. The total cost is calculated as the sum of data points Euclidean distances from their corresponding considered medoids in each iteration. This procedure is repeated until no further changes take place.

3.1. Determining optimal number of clusters

For estimating the validity of different numbers of clusters, a series of tests have been tried, and the number of clusters which, for most methods, was reported, was chosen.

The Silhouette validation technique [15] compares the average silhouette width of each cluster with that of the whole set of available data. In this way, a measure of tightness and separability (overall average silhouette) is defined, taking values from -1 to 1. The largest this parameter is, the better the clustering.

Calinski-Harabasz Index [16] measures intercluster over intracluster dissimilarity, using the between- and within-cluster scatter matrices as measures of dissimilarity. The most correct partition of the data, consequently, is supposed to occur for maximum values of Calinski-Harabasz index.

The Weighted inter-intra measure [17] is an index of homogeneity of the data to the separation with k clusters.

Krzanowski and Lai Index [18] is defined as

$$KL(k) = \frac{Diff(k)}{Diff(k-1)} \tag{1}$$

With $Diff(k) = (k-1)^{2/p} W_{k-1} - k^{2/p} W_k p$. p stands for the feature dimensionality, while W_k is the within cluster sums of squares of the partition. The optimum number of clusters is obtained for the maximum value of $KL(k)$

Davies-Bouldin Validity Index [19] measures the ratio between within-cluster scatter and between-cluster separation. It is defined as

$$DB(k) = \left(\frac{1}{k}\right) \sum_{i=1}^{k} \max_{i \neq j} \left\{ \frac{D_k(c_i) + D_k(c_j)}{D_k(c_i, c_j)} \right\} \tag{3}$$

k, D_k and $D(c_i, c_j)$ being the number of clusters, average distance of all data of a cluster to their cluster center c_i and distance between cluster centers, respectively. The optimal number of clusters is obtained for small values of index DB.

4. Experimental results

Experimenting for all 58 players, using as input Hav and H_l , as well as average time (T_{av}) each player can play a game without losing, most of the above tests reported k=3 as the optimal number of clusters (Fig. 5). The resulted clustering (Fig. 6) shows that one of the three classes (cluster 1) corresponds to moderate values in Hav and H_l and moderate to high play time. Cluster 2 corresponds to low head expressivity and low to moderate playtimes, while cluster 3 corresponds to large head expressivity and low amounts of average time per play for the corresponding players.

Statistical analysis shows that age and familiarity with game playing have strong links to the classes derived by the PAM algorithm. More in particular, older players (avg 33.9 yrs old, std 6.3 yrs) mostly belong to cluster 3 (high expressivity, low amounts of playtime), while not much difference exists between clusters 1 and

2, where players would manage to play for longer times, with lower head expressivity than cluster 3. It is typical that all players belonging to cluster 1, reported, at least a little experience in Super Mario, explaining, in this way, higher times achieved. In terms of gender, not a lot of significance occurred for the three classes, although, in cluster 3, 11 out of 17 players were women (large head expressivity and low amounts of average time per play). Allocation in the other clusters was not of high significance.

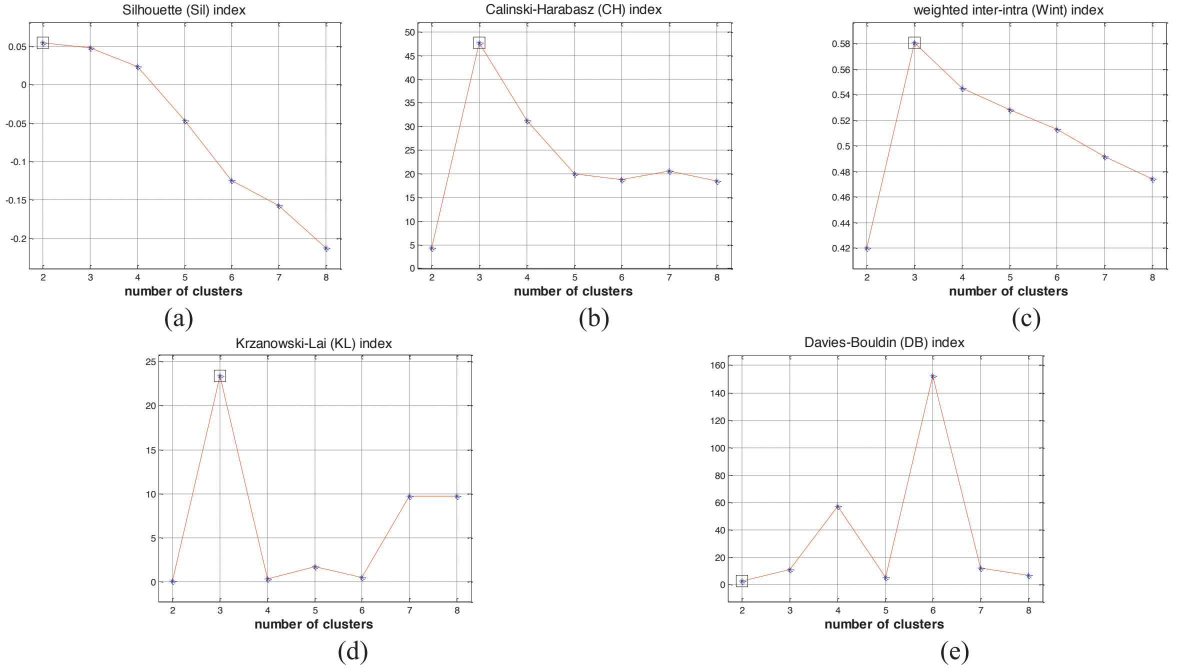


Fig. 5. Validity tests for optimal number of clusters with (a) Silhouette validation technique; (b) Calinski-Harabasz Index; (c) The Weighted inter-intra measure; (d) KL and (e) Davies-Bouldin Index

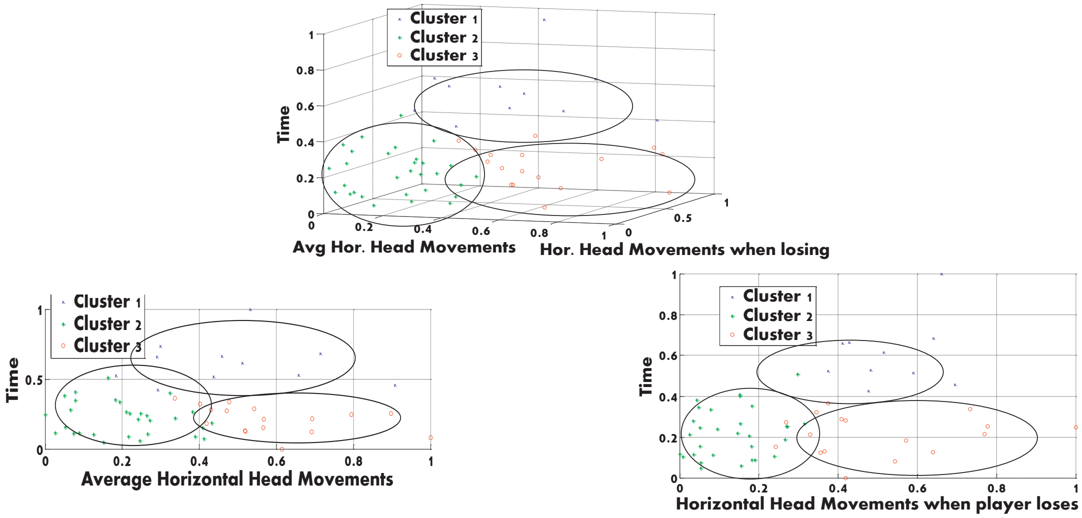


Fig. 6. Player Classification according to average play time and head movement characteristics

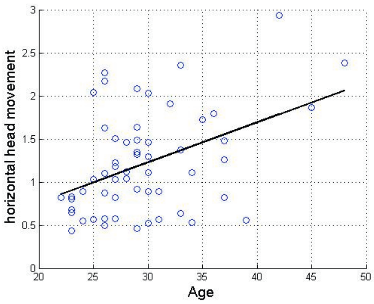


Fig. 7. Lip movements of male participants, as functions of age, for the events of losing.

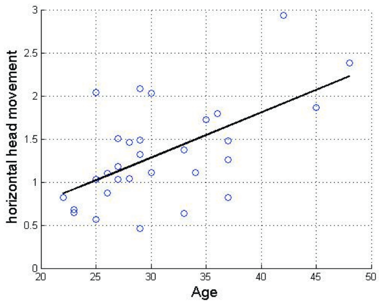


Fig. 8. Eye gaze movements, for female particpants, as functions of age, for the events of losing

Another significant criterion was that of general game play (p=0.002). In particular, in cluster 3, 65% of the players reported they do not play games at all, in contrast to Cluster 2, where 83.3% of the players said that they do play games, at least occasionally.

5. Discussion and conclusions

In this work, we have presented findings regarding player clustering during gameplay. We have utilized data coming from visual information (overall head expressivity and expressivity during critical events) as well as indicates of player performance. Further features have also appeared to have correlations with player characteristics (for instance, small facial gestures like lip biting or talking during critical events were more frequently met among male players and were a good indicator of expressiveness, while eye gazing was highly met among all participants (especially females) when Mario was killed) but did not prove sufficient as inputs for clustering, as they show mostly tendencies, instead of generic features. Figures 7 (r=0.46, R^2 =0.21, F=6.8, p=0.015) and 8 (r=0.55, R^2=0.30, F =12.07, p=0.0017) show the corresponding regression models for mouth activity for men and eye gaze for women, respectively. Lip movements where associated with changes at the saturation values on a frame-by-frame basis, while eye movements where associated with eye centers' position with regards to facial vertical boundaries. To measure the overall activation of this measurement, the first derivative was calculated.

Results have shown us that there exist links among expressivity, performance, experience and age. Near future research will take into account a series of further factors, like more detailed game actions and play styles, and explore options of more analytical, hierarchical clustering schemes. Utilizing the notions of in-game actions, visual behavior and player profile in a user modeling scheme, game or gamified applications can build on personalized content, targeting certain groups of users and, on a higher level, specialize content according to personal characteristics and needs. The above issues can play a very important role in creating installations such as games and learning environments, able to distribute personalized content and, in this way, maximize user engagement and flow. Although the presented work deals with the simple platform game of Super Mario, these ideas can adapt or expand to other game applications. Serious Games are designed to maximize engagement and immersion: Letting the AI component of a Serious Game decide regarding expected or unexpected behaviors could trigger certain events or could drive certain adaptation mechanisms. Linking visual and context-aware systems, as well as expected/unexpected behavior to certain levels of different cognitive/emotional states can give boost to further enhancements in the area of games designed for certain (educational, training, etc.) applications.

Acknowledgements

This research was supported by the FP7 ICT project SIREN (project no: 258453).

References

[1] D. W. Kolar, D. C. Funder, and C. R. Colvin, Comparing the accuracy of personality judgements by the self and knowledgeable others. *Journal of Personality 1996*, vol **64**(2): pp. 311–337

[2] N. Ducheneaut, N. Yee, E. Nickell, and R. J. Moore, "alone together?": exploring the social dynamics of massively multiplayer online games," Conference on Human Factors in computing systems, ser. CHI '06. New York, NY, USA: ACM, 2006, p. 407–416.

[3] Y. Matsumoto and R. Thawonmas, "Mmog player classification using hidden markov models," Third International Conference on Entertainment Computing - ICEC 2004, Eindoven, Netherlands, vol. 3166. Springer, 2004, p. 429–434.

[4] A. Tveit, O. Rein, J. V. Iversen, and M. Matskin, "Scalable agent-based simulation of players in massively multiplayer online games," 8th Scandinavian Conference on Artificial Intelligence, Frontiers in Artificial Intelligence and Applications. IOS. Press, 2003.

[5] N. Yee, N. Ducheneaut, H.-T. Shiao, and L. Nelson, "Through the azerothian looking glass: mapping in-game preferences to real world demographics," CHI, 2012, pp. 2811–2814

[6] G. van Lankveld, P. Spronck, H. J. van den Herik, and A. Arntz, "Games as personality profiling tools," IEEE Conference on Computational Intelligence and Games, 2011, pp. 197–202.

[7] D. Ramirez-cano, S. Colton, and R. Baumgarten, "Player classification using a meta-clustering approach," 3rd Annual International Conference Computer Games Multimedia Allied Technology 2010, pp. 297 – 304.

[8] A. Drachen, L. E. Nacke, G. Yannakakis, and A. L. Pedersen, "Correla-tion between heart rate, electrodermal activity and player experience in first-person shooter games," 5th ACM SIGGRAPH Symposium on Video Games, ser. Sandbox '10. New York, NY, USA: ACM, 2010, pp. 49–54.

[9] S. Asteriadis, N. Shaker, K. Karpouzis, and G. N. Yannakakis, "To-wards players affective and behavioral visual cues as drives to game adaptation," LREC Workshop on Multimodal Corpora for Machine Learning, Istanbul, 2012.

[10] N. Shaker, S. Asteriadis, G. Yannakakis, and K. Karpouzis, "A game-based corpus for analysing the interplay between game context and player experience," in EmoGames workshop, International Conference on Affective Computing and Intelligent Interaction (ACII2011), October 9, Memphis, USA., 2011.

[11] P. A. Viola and M. J. Jones, "Rapid object detection using a boosted cascade of simple features," Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, 2001, pp. 511–518.

[12] S. Asteriadis, N. Nikolaidis, and I. Pitas, "Facial feature detection using distance vector fields," *Pattern Recognition* 2009, **42**(7): 1388– 1398.

[13] J. Kovac, P. Peer, and F. Solina, "Human skin colour clustering for face detection," IEEE International Conference on Computer as a Tool, vol. 2, 2003, pp. 144–148.

[14] P. Rousseeuw and L. Kaufman, "Clustering by means of medoids," Statistical Data Analysis Based on the L1 Norm and Related Methods, pp. 405–416, 1987.

[15] P. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics* 1987, vol. 20, p. 53–65.

[16] R. B. Calinski and J. Harabasz, "A dendrite method for cluster analysis" *Communications in Statistics* 1974, **3:**1–27.

[17] A. Strehl, "Relationship-based clustering and cluster ensembles for high-dimensional data mining," Ph.D. dissertation, The University of Texas at Austin, 2002.

[18] W. Krzanowskiand and Y. Lai, "A criterion for determining the number of groups in a dataset using sum of squares clustering," *Biometrics* 1985, **44**:23–34.

[19] D. L. Davies and D. W. Bouldin, "A cluster separation measure*," IEEE Trans. Pattern Anal. Machine Intell* 1979, **1**(4): 224–27.