SEEV-Effort - Is it Enough to Model Human Attentional Behavior in Public Display Settings?

Benedikt Gollan

Pervasive Computing Applications
Research Studios Austria FG,
Thurngasse 8/20, 1090, Vienna, Austria
Email: benedikt.gollan@researchstudio.at

Alois Ferscha

Institute for Pervasive Computing
Johannes Kepler University,
Altenberger Strae 69, 4040 Linz
Email: ferscha@pervasive.jku.at

***Abstract*—Due to ever-increasing information overload, human attention has evolved to the most critical parameter in the successful design of pervasive display systems. This article aims at the validation of physical effort as a suitable descriptor for attentional and perceptional behavior in interactive public display scenarios. For this purpose, we integrated our qualitative effort-based behavior description approach into Wickens' established Saliency-Effort-Expectancy-Value (SEEV) attention model to compare predicted and observed attentional behavior and demonstrate its significance for attention mechanisms. The SEEV attention model is adapted to a public display scenario, with analysis focused on Minimum Required Effort (MRE) for the assessment of information. This attention modeling approach is evaluated based on data collected in an empiric study in an exhibition setting and is based on a database of 188 visitors. We employ different approaches of correlation analysis to evaluate the dependencies between required effort for information assessment and overt attention behavior which results in a maximum overall correlation score of 0.749 in a frame by frame correlation analysis.**

***Keywords*–attention; behavior analysis; public displays; implicit interaction**

I. Introduction

Human attention has revealed as being among the most crucial, yet least understood design criteria for modern, successful Information and Communication Technologies (ICT) systems. Digital information society technologies like the WWW, social networks or mail and messaging systems continuously wash floods of information to individuals via personal (mobile computers, smartphones) and public ICT systems (public displays, digital signage) [1], making it difficult for the individual to allocate attention to the right things at the right time and filter out the gold nuggets of information, which are relevant in the respective situation and context.

Given the overabundance of information approaching individuals, it has to be of immediate interest to include a profound understanding of (i) how attention is allocated and how it can be, (ii) if not explicitly measured, at least be estimated and modeled into technical solutions in public display scenarios. Only a fundamental understanding and successful modeling of human attention processes will allow the development of pervasive ICT systems that are designed for the benefit of users, and not only aiming at a most efficient distribution of arbitrary information to a maximum audience.

In the following, the current state of the art will be addressed. In Section 2, the underlying SEEV attention model and necessary simplifications for this work are described. In Section 3, the underlying experiment is introduced based on which the behavior data computation is carried out (Section 4). Results are presented and discussed in Section 5, to be concluded a discussion in Section 6.

A. Related Work

Attention research has evolved from first general observations by James [2] in the 1880s, over different Single Channel Theories [3–7] in which mental processes are regarded as serial competitive activities, over Capacity Theories in which human attention is not limited by perception bandwidth or number of channels but by processing capacity [8–10] to nowadays multi-channel and multiple-resource theories [11, 12]. These theories include multitasking capabilities and at the same time investigate distinct aspects of attention in detail. All these single models contribute to the overall understanding of the complex matter of human attention, yet, the latest generation of multiple-resource attention models provide promising capabilities of broad applicability in real technical applications.

Aiming at estimating attention levels using pervasive computing technologies, we depend on overt, observable somatic indicators of human attention which can be measured, quantified and interpreted. Concerning the individual, the analysis of Visual Attention represents the most intuitive and most frequently pursued approach [13], focusing mainly on stimulus-driven (bottom-up) attention mechanisms in contrast to expectation-driven (top-down) aspects of attention control. Modeling of Visual Attention is widely based on visual saliency as characteristics describing attention capturing capabilities of input stimuli. There are numerous models mimicking biological findings to best possible reproduce human eye movement patterns as, e.g., cognitive models [14][15], Bayesian models [16–18], decision theoretic models [19], information theoretic models [20][21] or alternatively pure data-driven models [22].

In technical applications, the so-called *Visual Focus of Attention* (VFOA) has become the main representation of the estimated orientation of attention, covering analysis of gaze direction [23], eye movements [24, 25], saccades and fixations [26, 27] or head orientation depending on sensor data quality. In addition to visual attention, overt attentional behavior is interpreted in public area scenarios, describing competitive

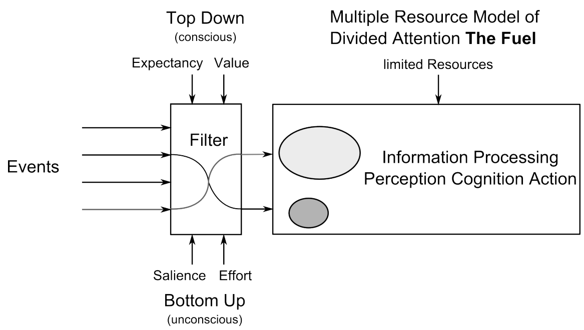

Figure 1. SEEV-Attention Model by Wickens [11] in which perception filters are controlled via Salience, Expectation, Effort and Value.

selection and switched attention processes. Smith et al. [28] created a head tracking in an out-of-home advertisement scenario in which subjects were evaluated as *focused* if their gaze was directed at the display longer than a certain threshold. Leykin and Hammoud [29] used single camera surveillance footage to estimate the area and direction of interest of pedestrians with head pose and movement direction as features. Ozturk et al. [30] investigated the orientation of people in indoor environments, interpreting orientation as indicator for visual focus of attention. Yakiyama et al. [31, 32] estimated attention levels towards target objects on the basis of computed distance, orientation and movement speed. Quek et al. [33] analyzed feet positioning and movement orientation to estimate the the orientation of perception.

In this paper, we will explore and validate physical effort as a crucial and expressive indicator of attentional behavior in the context of public display scenarios. For this purpose, we employ predictions of attentional behavior derived from Wickens' SEEV attention model, which are based on a previously published, general effort-based behavior description [34]. We analyze behavior to calculate physical effort scores that need to be invested for the potential perception of displayed information to identify significant correlations between physical effort, behavior and resulting engagement and attention. The identified correlations are supposed to validate physical effort as a reliable indicator for human attention estimation in public display scenarios.

II. EXPLORATION OF THE SEEV ATTENTION MODEL

In his elaborate research, Wickens created an extensive attention model [11], [35], which describes attention as an information processing system with multiple channels and resource types, which is equipped with a filtering unit driven by conscious (top-down) and unconscious (bottom-up) processes (cf. Figure 1). The central message of this approach is, that parallel cognitive processes are possible without interference as long as they differ in occupied channels and required types of resources (e.g., visual tasks can be executed in parallel to auditory tasks).

In its computational annotation, the so-called SEEV-attention model describes the probability to attend an area of interest via a linear equation, which is based on four components and their respective scaling factors - namely, the contributive factors of stimulus attractiveness (Saliency **S**), expected information entropy (Expectancy **E**x) and importance of the accustomed information (Value **V**) in contrast to the physical and mental Effort **E**f required for the assessment of the respective information (cf. Equ. 1):

$$P(A) = s \cdot \mathbf{S} - ef \cdot \mathbf{E}f + (ex \cdot \mathbf{E}x + v \cdot \mathbf{V}) \qquad (1)$$

Due to its linear and intuitive design, the SEEV model has successfully found application in several different scenarios as (i) aviation: modeling of attentional behavior and attention errors of pilots and air traffic controllers [36–38] (ii) automotive sector: simulation of car driver behavior [39–42] as well as information and infotainment systems in the automotive environment and mobile service interfaces [43] (iii) general computer interaction research [44] as well as (iv) perception and reaction to alerts and notifications [45]. The listed approaches use eye tracking or other sensorial data for behavior input and compare predicted and observed behavior to evaluate the applicability of the SEEV model and the respective author's developments, whereas, depending on the available data and given use-case, the SEEV-equation were simplified or generalizing assumptions were made.

A. The SEEV model in a public display scenario

The control parameters of the SEEV model Figure 1) require some more elaboration and interpretation regarding our specific use-case of pervasive display scenarios:

- **Value** represents the personal evaluation of the relevance of events, stimuli or perceived information. It is based on personal experiences and preferences, thus together with expectancy describes the individual intrinsic components of the SEEV attention model.
  As these intrinsic parameters are not accessible in the given experiment (anonymous users, no preferences or previous behavior history available), this parameter is deprived from our knowledge and control. As the presented use-case at an annual sports fair represents an event with voluntary presence, we are confident to assume positive overall associations of the audience towards the event in general and to simplify personal value to a constant positive average score.
- **Expectancy** describes the expected information bandwidth or entropy of an object, event or area of interest. Expectancy is closely related to personal experience of information consumption and especially thus linked to Value. Areas of interest that provided useful, entertaining or helpful information in the past are more likely to be attended and frequented in the future [46].
  Large-scale displays have become a fundamental part of public life and represent established omnipresent and continuous sources of information. They have evolved towards an essential role in the public domain with an associated high bandwidth of information, regardless of the associated personal relevance of displayed information. Again, this gives us the confidence to reduce model complexity by setting the expectancy score to a constant positive value.
- **Effort** represents the physical or mental effort, necessary to change current behavior to attend areas of interest and assess the available information. For example, this includes overcoming physical distances as well as exhausting cognitive filtering processes in noisy environments. Effort represents the only inhibitor of attention allocation in the SEEV attention model.

In our preliminary work [34], [47] we used invested physical effort to interpret engagement and attention to public displays. In this work, we try to validate these findings by employing the SEEV model in which effort - as the only inhibitory factor - plays a crucial role in the assessment of information. Hence, the required effort for content perception will serve as the main variable to be observed and evaluated.

- **Saliency** expresses the attractiveness of an event or stimulus mainly regarding bottom up processes, often affected significantly by contrast. Salient events differ from their environment via a single or several perceptional characteristics (e.g., color, size, movement, volume, frequency range, etc.). An overload with irrelevant stimuli is reported to reduce performance [48] and highly salient stimuli are likely to capture attention even when irrelevant to the task [49].
  Besides Effort, Saliency represents the second crucial factor in this setting. In our approach, we try to assess to which extent effort alone is sufficient to describe observed attention behavior. Due to that goal, we employed highly salient video content displaying extreme sports performances to ensure high and constant attractiveness. As a consequence, we again assume Saliency as a constant score in favor of a direct sensitivity to effort as the variable of interest, in the awareness of the over-simplification of the model and the need to take this reduction of complexity into consideration in the discussion and evaluation process.

The consequences of the described assumptions and simplifications are presented in the adapted equation for SEEV computation:

$$P_A(t) = (s \cdot \mathbf{S} + ex \cdot \mathbf{E}x + v \cdot \mathbf{V}) - ef \cdot \mathbf{E}f(t) \quad (2)$$
$$= (C) - ef \cdot \mathbf{E}f(t) \quad (3)$$
$$= 1 - ef \cdot \mathbf{E}f(t) \quad (4)$$

Value, Expectancy and Saliency are combined to a single constant value and in the context of a description of probability, to the maximum probability of 1. This results in an indirect proportional relation between attentional behavior and required effort, which will be experimentally verified in the following. The exact value of C is not relevant in the following evaluation via correlation since constant values do not have any influence on correlation scores. Since we only want to investigate the dependency of attention from effort behavior, it is an intuitive step to set C to 1 as it decreases potential misinterpretations.

III. Experiment Description

To find significant correlations between attention distributions predicted by SEEV probabilities and observed and annotated attentional behavior, we employ experimental data, which already served as input for preliminary works in [47]. The experiment was set up at a public sports event in the scope of the Vienna City Marathon. A large-scale public display integrated in a standard City Light cabinet (1,86 x 1,30 m) equipped with two 50” monitors and a depth camera sensor was installed in the entrance area of the event, with the sensor covering the area in front of the display (cf. Figure 2) with an horizontal opening angle of 57 ° and a maximal depth range of 6 m. The display itself did not show any interactive behavior, but employed depth sensors were only used for data collection.

Figure 2. Setup of large-scale public display at Vienna Sports World Fair. Depth cameras were used to detect and extract behavior features for effort computation.

The recorded depth data was afterwards used to apply OpenNI [50] and Primesense [51] skeleton tracking algorithms and extract skeleton joint data, which was used to compute and analyze behavior [34].

IV. Behavior Analysis and Creation of Data set

A. Computation of Minimum Required Effort - MRE

In earlier works, we interpreted the overall amount of invested effort to deduce the quality of engagement with a public display (Directed Effort [34]). A person in a public space has the following independent degrees of freedom regarding general movement in the horizontal plane (i) movement speed v (ii) movement direction φ (iii) body orientation. Similar to earlier computations of effort we stick to the strict additive separation of these behavioral dimensions and define thresholds for possible perception for each single parameter. We employ the following annotation in the ongoing: $i \in v, \sigma, \varphi, \tau$ as placeholder for the respective behavioral dimensions, as well as i_T as representation of the respective perceptional threshold.

As the analysis of the SEEV model is not directed at the quality but at the probability of attention, we are now aiming for the effort threshold that needs to be invested to enable the perception of the displayed content (MRE). Hence, some adaptations to the computation of invested effort had to be made in comparison to earlier processes: Body orientation is split up into head orientation σ and upper body orientation τ to better model head turns in relation to the overall body orientation.

The computation of all parameters is based on the ratio of the difference between current behavior and the respective thresholds i_T to the overall possible change in this specific parameter dimension (cf. Equ. 5-9). This provides a percental rating to what degree the person needs to change its activities in the respective behavioral dimension to enable the perception of displayed content. If possible, the selection of these thresholds was based on scientific research, otherwise they have been set according to observations and the personal experience of the authors. The respective maximum reference scores have been set via empiric analysis of the overall data set.

B. Speed

Movement speed is characterized as the magnitude of the movement vector $v(t) = |\overrightarrow{v(t)}|$. The speed threshold has been set to a score, which resembles a medium walking pace in the authors implementations. In figure 3(a), this is visualized via a

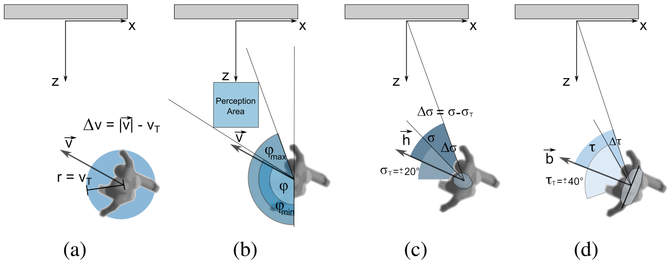

Figure 3. Computation of Minimum Required Effort in the dimensions of (a) movement speed $v(t)$, (b) movement direction $\varphi(t)$, (c) head orientation $\sigma(t)$ and (d) upper body orientation $\tau(t)$.

centric circle with radius v_T around the person center of mass location.

C. Movement Direction

According to the Focus-Nimbus Aspects Attributes and Resources (FN-AAR) model [52], which is based the Focus/Nimbus (FN) awareness model, there are areas of comfortable perception of displayed information. We employed the area as visualized in figure 3(b) as reference location for the perception threshold and to evaluate movement direction to overcome this hindrance. The position and movement vector of each proband is analyzed to compute whether the person is moving towards this perception zone, and if not, the required effort for behavior change (movement angle) is computed.

$$E(t) = e_v \cdot E_v(t) + e_\varphi \cdot E_\varphi(t) + e_\sigma \cdot E_\sigma(t) + e_\tau \cdot E_\tau(t) \tag{5}$$

$$E_v(t) = \begin{cases} \frac{v(t)-v_T}{\Delta v_{max}}, & \text{if } v(t) > v_T \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

$$E_\varphi(t) = \begin{cases} \frac{\varphi(t)-\varphi_T}{\Delta \varphi_{max}}, & \text{if } \varphi_{max} > \varphi(t) > \varphi_{min} \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

$$E_\sigma(t) = \begin{cases} \frac{\sigma(t)-\sigma_T}{\Delta \sigma_{max}}, & \text{if } \sigma(t) > \sigma_T \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

$$E_\tau(t) = \begin{cases} \frac{\tau(t)-\tau_T}{\Delta \tau_{max}}, & \text{if } \tau(t) > \tau_T \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

D. Head Orientation

Head orientation is interpreted as general Visual Focus of Attention as a low-level feature of eye gaze and description of general field of vision. The captured depth data did not provide sufficient resolution for actual eye tracking algorithms over the distance of several meters, so we were restricted to head orientation measurements derived from depth data. For this purpose, we employed the implementations by Fanelli et al. [53]. The threshold σ_t for head orientation was set to $20°$, as this angle between current eye position and goal of the saccade has been reported by Kahneman [8] as the limit from which head movements replaces eye movements for the assessment of the information (cf. Figure 3(c)). Below $20°$ head orientation in relation to the display location, we assume a perception of the display content as sufficiently probable.

E. Upper Body Orientation

The threshold for required turn of the upper body has been set to $40°$ (cf. Figure 3(d)). This seemingly arbitrarily defined score is the result of empiric observations and personal experiences and may be necessary to be replaced by more scientifically founded studies.

F. Ground Truth Labeling

Due to the lack of qualitative attention metrics, the evaluation of attention models represents a challenging issue. The main challenge in this context is obtaining an objective ground truth labeling of observed behavior. Subjective reports from subjects provide helpful assistance, yet in our case could not provide the required temporal resolution and detail.

In this work, we apply manually transcriptions that have already been applied in earlier works, which are combined from (i) real-time observations and annotations that were obtained in the gathering of the experiment data, (ii) detailed frame-wise post-hoc labeling of behavior scores and (iii) interview results from subjects regarding their overall awareness of the display and perceived content. The hand-labeled behavior scores are based on detailed categories of behavior that are associated to a numeric scale (0-10) and associated to 5 more general classes of behavior, representing increasing behavior change and engagement with the pervasive display.

In the following, we employ the general five behavioral classes of (0) No perception - display of of field of view, (1) Selective Perception - display in peripheral visual range, (2) Switched Attention - display actively in wandering gaze area, (3) Focused Attention - conscious perception, distinct behavior change (4) Sustained Attention - lasting perception, fundamental behavior change. Please refer to [47] for the detailed behavior assignment for the 11-tier scale and 5-class behavior segmentation.

G. Overall dataset

The resulting dataset is based on 188 probands that were observed and analyzed, resulting in an overall of 27'891 frames, which hold required effort for the respective behavioral dimension, prediction of attention probability and labeled ground truth of observed attentional behavior.

TABLE I. CORRELATION SCORES OF SINGLE EFFORT DIMENSIONS

Effort Parameter	Correlation Score
E_v	0.7138437
E_φ	0.3944934
E_σ	0.4232690
E_τ	0.2348277

V. Results and Discussion

As a first step, we tried to identify the weighting parameters e_i (cf. Equ. 5) to find an optimal contribution of the single behavioral dimensions for the aspired modeling of actual attention behavior. For this purpose, we computed the correlations between the single effort components and the hand-labeled ground truth data to evaluate the contributivity of the single parameters. The single correlation scores, averaged over the complete dataset, are displayed in Table I.

As can be observed, movement speed shows by far the most promising correlation of the four parameters. The minor

Copyright (c) IARIA, 2016. ISBN: 978-1-61208-461-9

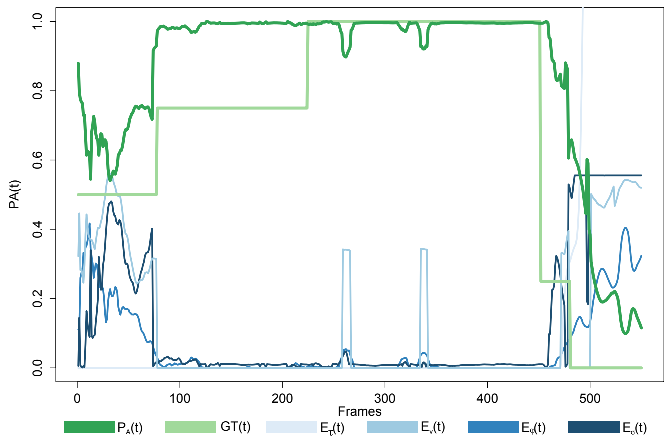

Figure 4. Exemplary sequence of behavioral dimensions E_v, E_φ, E_σ, E_τ, overall predicted attention probability $P_A(t) = 1 - E(t)$ and ground truth labels for a single proband over time (single correlation score: 0.84).

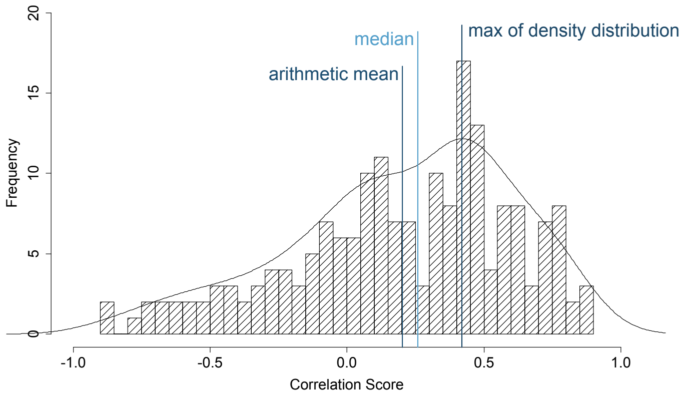

Figure 5. Histogram and density function (Gaussian kernel) of correlation results per person, with different statistical interpretative norms (mean, median, max. of density function).

performance of other parameters especially head and body orientation is partly related to the varying quality of skeleton tracking data. Yet these other behavioral aspects may still push the overall correlation between model prediction and observation scores. For this purpose, we used the correlation scores as weighting parameters:

$$e_i = \frac{cor(E_i)}{\sum_i cor(E_i)} \qquad (10)$$

A. Evaluation

For the analysis of the collected dataset, we employed the implementation of Pearson's product-moment correlation in R [54]. An exemplary plot of the sequence of computed scores for a sample of the dataset is displayed in Figure 4. It shows how the single computed parameters add up to the overall Effort score and relate to annotated behavior.

The overall evaluation of the gathered dataset using correlative dependencies requires some discussion. Generally, the computation of an adequate overall correlation score from different single correlation results is not trivial, as the database items differ in sample sizes and significance scores. The sample sizes range from 14 to 2321 captured frames per person with an average of 148.35 frames. This imbalance between data samples represents the main challenge regarding the computation of an expressive overall correlation score.

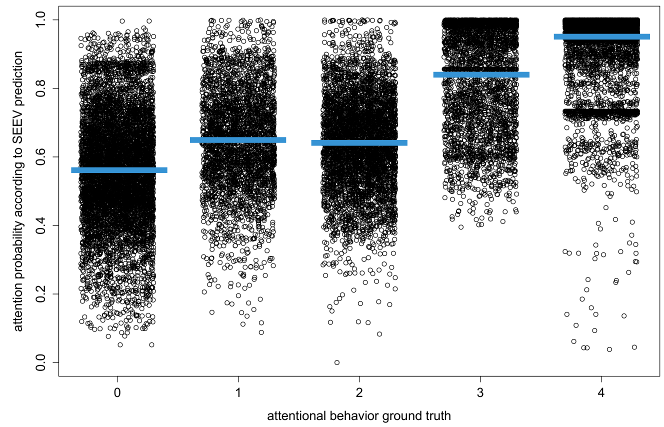

Figure 6. Scatter plot of computed attention probability scores for complete dataset (27'891) frames in relation to associated attention behavior ground truth. Horizontal bars indicate the arithmetic mean. The overall distributions show strong proportional tendencies.

To approach this issue of evaluation, we pursue two different evaluation approaches, neither of which can claim to give an absolute overall correlation score, but which will show the range of correlation dependencies and give an impression on the overall performance of our approach.

1) Averaging over single probands: Simple averaging over single subject correlation scores as overall correlation result based will show bias towards shorter sample sizes. As longer samples show a better correlation score, this process will underestimate the actual system performance. Furthermore, it is debatable, which averaging mechanism best represents correlation results. The histogram of obtained single correlation scores is visualized in Figure 5. In this case, the arithmetic mean seems to be too sensitive to far outliers, whereas the median (0.263) and the computed density function (0.413), employing a Gaussian kernel seem to better represent the overall distribution of results.

2) Creation of an overall Correlation: We can avoid the problem of averaging of incomparable samples via conflating all computed single temporal $P_A(t)$ and GT curves into a single, encompassing dataset. This would be equivalent to a single user being present in the scene for continuous 27'891 frames or 15h and 29min. The result is computed as the overall correlation between computed (effort-based) attention probabilities and hand-labeled observations, which computes to an overall correlation score of 0.749. This interpretation of correlation scores provides a far better rating of the dependencies between behavior and attention distribution, yet is probably prone to an overestimation since longer samples will cause heavier weights and again a imbalanced evaluation.

In the scatter plot in Figure 6, the general relation between predicted attention probabilities and overt attention behavior can be observed. The mean scores and general distributions show the expected strong proportional tendency, which is expected to increase under consideration of not only physical effort, but stimulus saliency as driving forces of attentional behavior.

Yet, this approach provides a new issue in the evaluation of these results in the assessment of significance. Usually, significance scores of $p < 0.05$ or $p < 0.02$ are claimed for reliable dependencies where coincidences are excluded. In the given evaluation, the problem arises from the extensive size of the dataset. As the significance score p is directly related to the size of the dataset it will turn against 0 ($2.2 \cdot e^{-16}$ in our case) for large data sets, a problem, which is already established in the scientific community. Thus the significance score loses its validity as an indicator in our case. Lin et al. [55] proposed to use confidence intervals instead of p as a more expressive evaluator regarding the validity of correlation results. In our case, the 95% confidence interval ranges from 0.743 to 0.754, a range, which makes us confident to report a significant relationship between required physical effort and attention behavior.

VI. Conclusion & Outlook

In this paper, we have identified and demonstrated significant correlations between attention behavior predicted by the established SEEV computational attention model, based on minimum required physical effort as descriptor, and actual observed attentional behavior in an empiric experimental study. In a critical discussion of different approaches towards a general evaluation of an absolute correlation score, we estimate the true correlation score between the overestimated score of 0.749 for overall correlation computation and the underestimated score of 0.413 for per person correlation averaging.

Taking into account, that the major aspect of time-dependent saliency has been ignored in the modeling and evaluation process, the obtained results absolutely show promising potentials regarding physical effort as a suitable attention modeling parameter in public domains. A future inclusion of real-time saliency evaluation of displayed visual and audio content is expected to further boost our attention modeling process. Additionally, investigating surplus effort, exceeding the necessary thresholds for perception might represent a suitable indicator for the quality of engagement and attention. Finally, in spite of undeniable efforts, the statistical evaluation requires further research regarding the computation of a suitable, explicit overall correlation score.

References

[1] T. H. Davenport and J. C. Beck, “The attention economy,” Ubiquity, vol. 2001, no. May, May 2001. [Online]. Available: http://doi.acm.org/10.1145/376625.376626

[2] W. James, The Principles of Psychology, Vol. 1. Dover Publications, Jun. 1950. [Online]. Available: http://www.worldcat.org/isbn/0486203816

[3] D. E. Broadbent, Perception and communication, 3rd ed. Oxford: Pergamon Press, 1969.

[4] K. J. Craik, “Theory of the human operator in control systems1,” British Journal of Psychology. General Section, vol. 38, no. 2, 1947, pp. 56–61.

[5] A. M. Treisman and G. Gelade, “A feature-integration theory of attention,” Cognitive Psychology, vol. 12, no. 1, 1980, pp. 97–136.

[6] J. A. Deutsch and D. Deutsch, “Attention: Some theoretical considerations,” Psychological review, vol. 70, 1963, pp. 80–90.

[7] N. Lavie, A. Hirst, J. W. d. Fockert, and E. Viding, “Load Theory of Selective Attention and Cognitive Control,” Journal of Experimental Psychology: General, vol. 133, no. 3, 2004, pp. 339–354.

[8] D. Kahneman, Attention and effort. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

[9] H. P. Bahrick and C. Shelly, “Time sharing as an index of automatization.” Journal of Experimental Psychology, vol. 56, no. 3, 1958, p. 288.

[10] W. Schneider and R. M. Shiffrin, “Controlled and automatic human information processing: I. Detection, search, and attention,” Psychological Review, vol. 84, no. 1, 1977, pp. 1–66.

[11] C. D. Wickens, “Multiple resources and mental workload.” Human Factors, vol. 50, no. 3, 2008, pp. 449–455. [Online]. Available: http://dblp.uni-trier.de/db/journals/hf/hf50.htmlWickens08a

[12] E. I. Knudsen, “Fundamental components of attention.” Annual Review of Neuroscience, vol. 30, no. 1, 2007, pp. 57–78. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/17417935

[13] A. Borji and L. Itti, “State-of-the-art in visual attention modeling,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 35, no. 1, 2013, pp. 185–207.

[14] L. Itti, C. Koch, and E. Niebur, “A model of saliency-based visual attention for rapid scene analysis,” IEEE Transactions on pattern analysis and machine intelligence, vol. 20, no. 11, 1998, pp. 1254–1259.

[15] O. Le Meur, P. Le Callet, D. Barba, and D. Thoreau, “A coherent computational approach to model bottom-up visual attention,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 28, no. 5, 2006, pp. 802–817.

[16] A. Oliva, A. Torralba, M. S. Castelhano, and J. M. Henderson, “Top-down control of visual attention in object detection,” in Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, vol. 1. IEEE, 2003, pp. I–253.

[17] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell, “Sun: A bayesian framework for saliency using natural statistics,” Journal of vision, vol. 8, no. 7, 2008, p. 32.

[18] J. Li, Y. Tian, T. Huang, and W. Gao, “Probabilistic multi-task learning for visual saliency estimation in video,” International Journal of Computer Vision, vol. 90, no. 2, 2010, pp. 150–165.

[19] D. Gao and N. Vasconcelos, “Discriminant saliency for visual recognition from cluttered scenes,” in Advances in neural information processing systems, 2004, pp. 481–488.

[20] R. Rosenholtz, “A simple saliency model predicts a number of motion popout phenomena,” Vision research, vol. 39, no. 19, 1999, pp. 3157–3163.

[21] N. Bruce and J. Tsotsos, “Saliency based on information maximization,” in Advances in neural information processing systems, 2005, pp. 155–162.

[22] X. Hou and L. Zhang, “Saliency detection: A spectral residual approach,” in Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on. IEEE, 2007, pp. 1–8.

[23] R. Stiefelhagen, M. Finke, J. Yang, and A. Waibel, “From gaze to focus of attention,” in Visual Information and Information Systems. Springer, 1999, pp. 765–772.

[24] E. Kowler, “Eye movements: The past 25years,” Vision research, vol. 51, no. 13, 2011, pp. 1457–1483.

[25] M. Hayhoe and D. Ballard, “Eye movements in natural behavior,” Trends in cognitive sciences, vol. 9, no. 4, 2005, pp. 188–194.

Copyright (c) IARIA, 2016. ISBN: 978-1-61208-461-9

[26] S. P. Liversedge and J. M. Findlay, "Saccadic eye movements and cognition," Trends in cognitive sciences, vol. 4, no. 1, 2000, pp. 6–14.

[27] J. E. Hoffman and B. Subramaniam, "The role of visual attention in saccadic eye movements," Perception & psychophysics, vol. 57, no. 6, 1995, pp. 787–795.

[28] K. C. Smith, S. O. Ba, J.-M. Odobez, and D. Gatica-Perez, "Tracking attention for multiple people: Wandering visual focus of attention estimation," IDIAP, Tech. Rep., 2006.

[29] A. Leykin and R. Hammoud, "Real-time estimation of human attention field in lwir and color surveillance videos," in Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on, june 2008, pp. 1 –6.

[30] O. Ozturk, T. Yamasaki, and K. Aizawa, "Tracking of humans and estimation of body/head orientation from top-view single camera for visual focus of attention analysis," in Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on, 27 2009-oct. 4 2009, pp. 1020 –1027.

[31] Y. Yakiyama, N. Thepvilojanapong, M. Iwai, O. Mihirogi, K. Umeda, and Y. Tobe, "Observing real-world attention by a laser scanner," IPSJ Online Transactions, vol. 2, 2009, pp. 93–106.

[32] N. Thepvilojanapong, Y. Yakiyama, O. Mihirogi, M. Iwai, K. Umeda, Y. Tobe, and R. Shibasaki, "Evaluating people's attention in the real world," in ICCAS-SICE, 2009, aug. 2009, pp. 3702 –3709.

[33] F. Quek, R. Ehrich, and T. Lockhart, "As go the feet...: on the estimation of attentional focus from stance," in Proceedings of the 10th international conference on Multimodal interfaces, ser. ICMI '08. New York, NY, USA: ACM, 2008, pp. 97–104. [Online]. Available: http://doi.acm.org/10.1145/1452392.1452412

[34] B. Gollan and A. Ferscha, "A generic effort-based behavior description for user engagement analysis," in Physiological Computing Systems. Springer, 2014, pp. 157–169.

[35] C. D. Wickens, "Multiple resources and mental workload," Human Factors: The Journal of the Human Factors and Ergonomics Society, vol. 50, no. 3, 2008, pp. 449–455.

[36] C. Wickens, J. McCarley, and K. Steelman-Allen, "Nt-seev: A model of attention capture and noticing on the flight deck," Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 53, no. 12, 2009, pp. 769–773. [Online]. Available: http://pro.sagepub.com/content/53/12/769.abstract

[37] C. D. Wickens, J. S. McCarley, A. L. Alexander, L. C. Thomas, M. Ambinder, and S. Zheng, "Attention-situation awareness (a-sa) model of pilot error," Human performance modeling in aviation, 2008, pp. 213–239.

[38] X. Wu, X. Wanyan, and D. Zhuang, "Attention allocation modeling under multifactor condition," Journal of Beijing University of Aeronautics and Astronautics, vol. 8, no. 39, 2013, p. 1086.

[39] W. J. Horrey, C. D. Wickens, and K. P. Consalus, "Modeling drivers' visual attention allocation while interacting with in-vehicle technologies." Journal of Experimental Psychology: Applied, vol. 12, no. 2, 2006, p. 67.

[40] J.-T. Wong and S.-H. Huang, "A microscopic driver attention allocation model," Advances in Transportation Studies an International Journal, 2011, pp. 53–64.

[41] R. Kaul, M. Baumann, and B. Wortelen, "The influence of predictability and frequency of events on the gaze behaviour while driving," in Human Modelling in Assisted Transportation. Springer, 2011, pp. 283–290.

[42] N. D. Cassavaugh, A. Bos, C. McDonald, P. Gunaratne, and R. W. Backs, "Assessment of the seev model to predict attention allocation at intersections during simulated driving," in 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, no. 52, 2013.

[43] J. Niemann, V. Presse, J. Reissland, and A. Naumann, "Developing a user-centered mobile service interface based on a cognitive model of attention allocation," in Human-Computer Interaction. Springer, 2010, pp. 50–57.

[44] B. Gore, B. Hooey, C. Wickens, and S. Scott-Nash, "A computational implementation of a human attention guiding mechanism in midas v5," in Digital Human Modeling, ser. Lecture Notes in Computer Science, V. Duffy, Ed. Springer Berlin Heidelberg, 2009, vol. 5620, pp. 237–246. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-02809-0_26

[45] J. B. Mulligan, "Factors influencing scanning for alerts," Journal of Vision, vol. 10, no. 15, 2010, pp. 69–69.

[46] J. Müller, D. Wilmsmann, J. Exeler, M. Buzeck, A. Schmidt, T. Jay, and A. Krüger, "Display blindness: The effect of expectations on attention towards digital signage," in Pervasive Computing. Springer, 2009, pp. 1–8.

[47] B. Gollan, B. Wally, and A. Ferscha, "Automatic human attention estimation in an interactive system based on behaviour analysis," Proc. EPIA 2011, 2011.

[48] W. Bacon and H. Egeth, "Overriding stimulus-driven attentional capture," Perception & Psychophysics, vol. 55, no. 5, 1994, pp. 485–496. [Online]. Available: http://dx.doi.org/10.3758/BF03205306

[49] S. Yantis, "Stimulus-driven attentional capture," Current Directions in Psychological Science, 1993, pp. 156–161.

[50] OpenNI User Guide, OpenNI organization, November 2010, last viewed 19-01-2011 11:32. [Online]. Available: http://www.openni.org/documentation

[51] Prime Sensor NITE 1.3 Algorithms notes, PrimeSense Inc., 2010, last viewed 19-01-2011 15:34. [Online]. Available: http://www.primesense.com

[52] S. Benford, A. Bullock, N. Cook, P. Harvey, R. Ingram, and O.-K. Lee, "From rooms to cyberspace: models of interaction in large virtual computer spaces," Interacting with Computers, vol. 5, no. 2, 1993, pp. 217–237.

[53] G. Fanelli, T. Weise, J. Gall, and L. Van Gool, "Real time head pose estimation from consumer depth cameras," in Pattern Recognition. Springer, 2011, pp. 101–110.

[54] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2012, ISBN 3-900051-07-0. [Online]. Available: http://www.R-project.org/

[55] M. Lin, H. C. Lucas Jr, and G. Shmueli, "Research commentary-too big to fail: large samples and the p-value problem," Information Systems Research, vol. 24, no. 4, 2013, pp. 906–917.