Slovak University of Technology in Bratislava

Faculty of Informatics and Information Technologies

Miroslav Laco

Dissertation Thesis Abstract

MODELLING VISUAL ATTENTION USING APPLIED PRINCIPLES OF COMPUTER VISION

to obtain the Academic Title of *philosophiae doctor (PhD.)*

Degree course:	Applied Informatics
Field of study:	Applied Informatics
Form of study:	Internal
Place of development:	Institute of Computer Engineering and Applied Informatics, FIIT STU Bratislava

Bratislava 2022

Dissertation Thesis has been developed at the Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava.

Submitter: Miroslav Laco
Institute of Computer Engineering and Applied Informatics
Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava

Supervisor: Prof. Vanda Benešová
Institute of Computer Engineering and Applied Informatics
Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava

Oponents: Assoc. Prof. Elena Šikudová
Department of Software and Computer Science Education
Faculty of Mathematics and Physics
Charles University in Prague

Assoc. Prof. Zdeněk Míkovec
Department of Computer Graphics and Interaction
Faculty of Electrical Engineering
Czech Technical University in Prague

Dissertation Thesis Abstract was sent: ....
Dissertation Thesis Defence will be held on .... at .... pm at the Institute of Computer Engineering and Applied Informatics, Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava (Ilkovičova 2, Bratislava).

Prof. Ing. Ivan Kotuliak, PhD.
Dean of FIIT STU in Bratislava

ABSTRAKT

Modelovanie l'udskej vizuálnej pozornosti je výskumnou doménou aplikovanej informatiky už niekol'ko desat'ročí. Ciel'om výskumu v danej oblasti je tvorba výpočtových modelov, ktoré simulujú proces l'udského vizuálneho vnímania. Takéto modely majú významné využitie v mnohých aplikačných oblastiach informatiky, a to od spracovania signálu až po súčasné, inovatívne algoritmy počitacového videnia a grafiky.

Tvorba výpočtových modelov vizuálnej pozornosti je netriviálnou úlohou pre komplexnost' procesu l'udského vizuálneho vnímania, ako aj pre náročnost' skúmania javov, ktoré vizuálnu pozornost' ovplyvňujú. V tejto dizertačnej práci sa zameriavame na najmodernejšie trendy v oblasti modelovania l'udskej vizuálnej pozornosti, vrátane metód aplikujúcich strojové učenie a umelú inteligenciu, s využitím rôznych architektúr neurónových sietí a personalizovaného učenia.

Odrážajúc sa od stavu súčasneho poznania máme za ciel' pokročilými výpočtovými algoritmami analyzovat' vizuálnu pozornost' človeka vrátane aspektov, ktoré ju ovplyvňujú, a tvorit' vlastné inovatívne modely vizuálnej pozornosti, ktoré sú zaujimavé pre d'alšie vedecké aplikácie v rôznych odvetviach. Chceme tiež uvažovat' doteraz nepreskúmané fenomény v modelovaní vizuálnej pozornosti, ktorými sú vnímanie aspektu hĺbky v rôznych typoch prostredia, vplyv vnútorného stavu človeka na vizuálnu pozornost', či personalizácia modelov vizuálnej pozornosti.

Kl'účové slová: počítačové videnie, modelovanie l'udskej vizuálnej pozornosti, vizuálna výraznost'

ABSTRACT

Visual attention modelling has been an open research domain in the field of applied informatics throughout the past decades. The ultimate aim of the research in the field is to propose computational models of visual attention which are able to simulate the complex process of human visual attention. These models are widely used in many application domains, including signal processing and novel computer vision and graphics algorithms.

It is a non-trivial task to formally define a computational model of human visual attention. The complexity of the visual attention processes is high, and there is a wide range of aspects affecting visual attention in a natural environment. We aim to focus on the novelties and recent trends in visual attention modelling, including innovative machine learning methods, artificial intelligence and personalised attention modelling in this dissertation thesis.

Based on the analysis of the state-of-the-art in visual attention modelling, we aim to analyse human visual attention alongside the aspects affecting attention using advanced computational algorithms and propose our own innovative models of visual attention with perspective for future applications. Our goal is to focus on the phenomena which were not thoroughly studied before. These include binocular depth perception in various types of the environment, the impact of the internal state of the observer on visual attention, or the recent research field of personalised attention modelling.

Keywords: computer vision, visual attention modelling, visual saliency

Contents

1 Introduction

The main principles of the attention mechanisms are prioritization and selectivity to choose the most significant stimuli from the perceived scene to process [22]. Visual attention is a set of non-trivial phases and processes and, thus, attracted researchers many decades ago to start its exploration. Nowadays, we know much about visual perception, and attention from the neuro-psychological point of view [21, 22, 66]. However, we are still not able to define visual attention completely, and further research is required to help us better understand this phenomenon. One of the biggest challenges related to visual attention research is the formal definition of attention along with building up visual attention models. Much effort has been put into visual attention modelling during the past decades [8, 7]. However, the models are still not able to simulate human attention behaviour under specific conditions and to treat it as a subject of individuality and uniqueness [7, 23].

The main contribution of this thesis can be divided into these points:

1. analysis and quantification of the exogenous bias of the depth cues on the binocular vision in different types of the environment (real environment, virtual environment, two-dimensional captures of the real environment) with a comparative analysis based on the self-proposed empirical study with artificial scene setup using visually neutral objects placed in the peripheral visual field in various distances from the observer as proposed by Wang et al. [70],
2. analysis and quantification of the endogenous bias of the internal state of the observer on the visual attention in the means of the affective, emotional state based on the self-proposed empirical study,
3. methodology for personalization of the visual attention models as a personalized attention fingerprint of the individuals based on the endogenous bias of the individuality on the visual attention for generating personalized attention predictions over unseen visual stimuli.

2 Visual Perception

Eye movement analysis is essential for measuring visual attention and creating datasets containing information about people's visual attention.

Fixation is a period during which our visual system collects visual information from a specific part of the scene, analyses the collected information, and forms the decision (i.e. reaction to the stimulus) [39, 62].

Saccade refers to a period during which the eye is physically directed by fast moving towards the very next stimulus at the scene [39, 62].

The aspects which influence and drive attention are generally divided into two main groups based on their characteristics. **Bottom-up aspects** are scene-driven, which means that their origin is in the perceived scene or image itself [8, 7]. Some bottom-up aspects (i.e. colour, intensity, shape) can be perceived from a single retinal image; hence, they are subject to monocular vision. There is a specific group of aspects that are bound to a combination of monocular and stereoscopic binocular vision and a three-dimensional environment [37]. A combination of bottom-up aspects is a base for the so-called saliency of a particular part of the scene. Bottom-up saliency is a measure of the significance of parts of the scene that drives the attention [8]. The more salient a part of the scene is, the higher the probability of grabbing our attention. In contrast, **top-down aspects** have their origin inside us and are driven by our prior knowledge, experience, memory, goals or visual task performed over the scene (also known as goal-driven aspects) [8].

The prerequisite for the visual perception process to be initiated is a visual scene containing a set of visual elements, which enters our visual system through the eyes as a visible wavelength spectrum. Visual elements and their interposition in the scene itself carry a certain level of bottom-up-based saliency. The saliency of these scene parts then compete for our attention resulting in producing a set of measurable fixations on certain parts of the visual scene in a particular order based on their importance [39, 62]. The importance is influenced not only by the bottom-up saliency of the scene parts but also by the top-down features. The context building up the top-down effect on the visual attention also includes the actual internal state of the individual, which consists of a stable background the individual carries (i.e. gender, education, previous experience, memories) and a temporal internal state of the observer which can change over time (i.e. emotional state).

3 Visual Attention Modelling

To model the complex process of visual attention, we first need to capture the data about the visual attention of a subject. Non-intrusive and the most common method for measuring attention is eye-tracking. The output of the eye-tracking is often a reference to the frame of the perceived scene and the point coordinates of the gaze position at the frame [17]. The measured visual attention data in the means of the time-series of gaze points representing the attention behaviour of the observers are usually classified into fixations and saccades using standard algorithms such as velocity threshold identification (I-VT) or dispersion threshold identification (I-DT) fixation classifier (more in Komogortsev et al. [34]). There are dozens of visual attention datasets from the past 20 years.

SALICON was introduced by Jiang et al. [33]. It consists of 20000 images of natural scenes taken from the MS COCO dataset [47]. The visual stimuli were viewed by 60 participants of the mouse tracking study.

CAT2000 was published by Borji et al. [9]. The dataset consists of 4000 visual stimuli classified into 20 categories based on the nature of the presented natural or artificial scenes. The authors guarantee that at least 18 statistically relevant participants viewed each stimulus.

Visual attention modelling is a research area that has its base in the early 1980s when Treisman and Gelade [69] formalized their feature-integration theory. dThe feature integration theory and the computational model proposal inspired the research group of Itti et al. [29] to build up one of the first computational saliency models, referred to as a baseline nowadays.

One of the largest groups of the attention models were cognitive models [28, 27], Bayesian models [75, 68, 27], decision theoretic models [20, 50], information theoretic models [11, 12, 26, 52] and graphical models [49, 2, 24]. A large group of the latest models are based upon Convolutional Neural Networks (CNNs), which have an outstanding ability to learn and predict salient regions from the multi-scale features using the typical multi-layer architecture [7]. One group of models based on the CNNs are models which extract high-level features from the images and apply the Multi-Layer Perceptrons as the last layer [76, 41, 71, 45]. The second group of models is based on the CNNs are models which apply a Fully-Convolutional Layer as the last layer of the network [25, 43, 44].

Standardized evaluation methods and metrics have developed throughout the last decades of visual attention modelling. These include Area under Receiver Operating Characteristic, Normalized Scanpath Saliency, Linear Pearson's Correlation Coefficient, Kullback-Leibler divergence, Similarity and Information Gain [13].

4 Research theses

4.1 Exogenous Bias of Depth Cues on Binocular Vision

Depth cues are essential aspects that bias the attention significantly under certain conditions [14, 61, 40, 31]. There is a lack of a robust comparative study on the depth bias of the attention in various environmental conditions, speaking about binocular depth cues perception:

- in the real environment,
- in the virtual reality,
- from a wide-screen display.

Based on the conclusions of Itti et al. [30] and Wang et al. [70], we formulate a null hypothesis and alternative hypothesis about the depth *pop-out* effect as follows:

- $H1.1_0$: Fixation density distribution on the identical objects in the given visual scene is equally distributed regardless of the differences in the distances of the objects from the observer in the scene from the neighbouring objects.
- $H1.1_1$: Fixation density distribution on the identical objects in the given visual scene is not equally distributed and is dependent on the differences in the distances of the objects from the observer in the scene from the neighbouring objects.

The independent variable for the proof of this hypothesis is the distance of one part of the visually neutral scene. The dependent variable, in this case, is the distribution of the fixation densities on the parts of the visually neutral scene.

We formulate a null hypothesis and alternative hypothesis about the effect of the distances of the parts of the visual scene from the observer on the saliency of these parts as follows:

- $H1.2_0$: Fixation density distribution on the identical objects in the given visual scene is equally distributed regardless of the distance of the objects from the observer in the scene.
- $H1.2_1$: Fixation density distribution on the identical objects in the given visual scene is not equally distributed and is dependent on the distance of the objects from the observer in the scene.

The independent variables for the proof of this hypothesis are the distances of the parts of the visually neutral scene. The dependent variable, in this case, is the distribution of the fixation densities on the parts of the visually neutral scene.

4.2 Endogenous Bias of Emotional State on Visual Attention

We expect that human visual attention behaves differently not only when being exposed to various types of stimuli but also for various manipulations with the emotional bias of the subjects [6, 1].

Based on the nature of our perspective of view on the emotionally affected attention bias, we pose a null hypothesis about the impact of distraction, engagement and motivation aspects and their affect on the visual attention of an observer in terms of visual search performance as follows:

- $H2.1_0$: Manipulating the original emotional state of a subject with a positive emotional bias has not an impact on one's visual search performance.
- $H2.1_1$: Manipulating the original emotional state of a subject with a positive emotional bias has an impact on one's visual search performance.

The independent variable for the proof of this hypothesis is the emotional bias of an observer (either positive or neutral) in comparison to the neutral emotional state of the subject by the start of the experimental study. The dependent variable, in this case, is the distribution of the results of subjects' visual search performance while accomplishing various visual search tasks.

Based on the nature of our perspective of view on the emotionally affected attention bias, we pose a null hypothesis about the impact of distraction, engagement and motivation aspects and their affect on the visual attention fingerprint of a subject in terms of distribution of fixations for a set of given visual stimuli as follows:

- $H2.2_0$: Manipulating the original emotional state of a subject with a positive emotional bias does not change one's distribution of fixations for a set of given visual stimuli.
- $H2.2_0$: Manipulating the original emotional state of a subject with a positive emotional bias do change one's distribution of fixations for a set of given visual stimuli.

The independent variable for the proof of this hypothesis is the emotional bias of an observer (either positive or neutral) in comparison to the neutral emotional state of the subject by the start of the experimental study. The dependent variable, in this case, is the distribution of fixations for a set of given visual stimuli.

4.3 Endogenous Bias of Individuality on Visual Attention

Nowadays, there is strong evidence that personalised attention research and modelling would help us not only to open new possibilities in application fields related to attention modelling [23, 35], but even enhance the precision of the existing models [42, 72, 46, 74].

Based on the nature of our assumptions, we pose a null hypothesis and alternative hypothesis about the proposed personalisation methodology of the generalised visual attention model as follows:

- $H3.1_0$: Personalisation methodology applied on the chosen generalised visual attention model does not result in a personalised visual attention model with higher accuracy of the visual attention predictions for a given set of visual stimuli for a concrete subject than the basal generalised visual attention model.
- $H3.1_1$: Personalisation methodology applied on the chosen generalised visual attention model results in a personalised visual attention model with higher accuracy of the visual attention predictions for a given set of visual stimuli for a concrete subject than the basal generalised visual attention model.

The independent variable for the proof of this hypothesis is the proposed personalisation methodology applied on the chosen generalised visual attention model. Dependent variables, in this case, are the results of standard metrics for evaluation of the visual attention models evaluated over a test dataset containing visual stimuli and the ground-truth fixation maps the models have never seen.

Based on the assumptions, we pose another null and alternative hypotheses as follows:

- $H3.2_0$: Distribution of the AUC-Judd improvements of the visual attention predictions generated by the personalised visual attention model over the basal generalised visual attention model evaluated on the statistically relevant number of subjects is from a normal distribution with the mean laying in the interval $< -0.005; +0.005 >$.
- $H3.2_1$: Distribution of the AUC-Judd improvements of the visual attention predictions generated by the personalised visual attention model over the basal generalised visual attention model evaluated on the statistically relevant number of subjects is from other than the normal distribution.

The independent variable for the proof of this hypothesis is the proposed personalisation methodology applied on the chosen generalised visual attention model. The dependent variable, in this case, is the distribution of the improvements of the visual attention predictions generated by the personalised visual attention model over the basal generalised visual attention model evaluated by the AUC-Judd metric on the statistically relevant number of subjects.

5 Exogenous Bias of Depth Cues on Binocular Vision

Research studies related to the depth perception and modelling of the depth cues are generally divided into two main categories:

- studies which examine monocular depth cues [63, 18, 51],
- studies which examine depth cues as a combination of the monocular and binocular cues [70, 57, 37].

A special category of studies is the one that aims to compare depth perception from the two-dimensional and three-dimensional environment and, hence, is a combination of both of the categories [14, 61, 40, 31].

During the following years, the utilization of stereoscopic displays in attention research greatly influenced the examination of depth cues. Stereoscopic displays were used in the studies of Lange et al., and Wang et al. [40, 70]. Both of the research teams projected static stereoscopic images to participants during the eye-tracking study. While in the work of Lang et al. [40] it was a set of 600 images from the natural environment, in the work of Wang et al. [70] it was a set of artificial images where all other aspects biasing the attention were eliminated. Lang et al. found out that depth cues bias the attention more significantly at farther locations when it comes to the real environment. They state that the relationship between the object's saliency and its distance from the observer is non-linear. Moreover, closer objects were fixated by participants of their study more frequently than the farther ones.

Wang et al. [70] addressed the problem of quantification of depth bias on visual attention. The quantification of such a complex phenomenon is a non-trivial task and requires isolation of the examined cues from the influence of the others. Therefore, they proposed an experimental study to explore the impact of depth cues on attention while viewing artificial stereoscopic images with the suppressed influence of other aspects influencing attention. The artificial scene was made up of regular salt and pepper noise in the background and 8 static identical white circular objects. Each object could appear at a different distance from the observer in the stereoscopic three-dimensional space. None of the objects was situated in the central visual field of the observer to suppress strong visual centre bias. Wang et al. found out that 20% of the total fixations were bound to the closest objects, them the most salient from the depth perception perspective of view [70]. However, the farthest objects were fixated significantly earlier than the closer ones.

We proposed and conducted extensive eye-tracking experimental studies focused on the visual depth bias to assess our hypotheses. We build up mainly on the work of Wang et al. [70] and use a similar visual scene setup as proposed in their work in a different kind of more realistic environment. Many improvements were made to the original proposal of Laco et al. [37] by Laco et al. [38]. The schematic view of the octagonal object configuration for the eye-tracking study, alongside the object labelling, is visualised in Figure 1 and Figure 2. The setup was arranged in a laboratory with a visually neutral background.

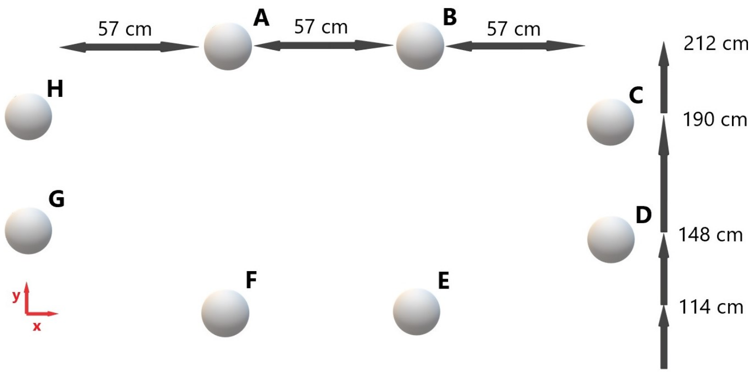


Figure 1: Schematic view of the proposed object configuration from the observer's perspective of view (frontal view). The objects are labelled and arranged in the octagonal configuration to avoid the strong visual centre-bias.

The eye-tracking study with the wide-screen display took place with 25 participants (aged $\overline{x} = 21.3; \sigma = 1.2$ years, where $\overline{x}$ further denotes mean value of the sample and σ denotes standard deviation), while the eye-tracking study in virtual reality took place with 38 participants (aged $\overline{x} = 22.1; \sigma = 0.7$ years). The acquired dataset from the eye-tracking studies in reality contained the eye-tracking data from 28 participants (aged $\overline{x} = 21.7; \sigma = 0.8$ years).

The first configuration type was evaluated with an aim to prove trends of higher fixation density on the *popping-out* objects to assess our null hypothesis $H1.1_0$. We found out statistically significant differences in the fixation density distributions and a weak positive correlation between fixation densities and the objects' depth *pop-out* effect on the visual saliency ($r = 0.145; p < .001$). We reject the null hypothesis $H1.1_0$ and accept the alternative hypothesis $H1.1_1$.

For other object configuration types, we found out statistically significant differences in the fixation density distributions using the ANOVA test. We found out that we have to reject $H1.2_0$ for all configurations for the virtual reality and real environment, thus, accepting alternate hypothesis $H1.2_1$ about the existence of the differences between fixations densities on the objects in different distances from the observer. Furthermore, we were looking for a positive correlation between the fixation densities and object depths

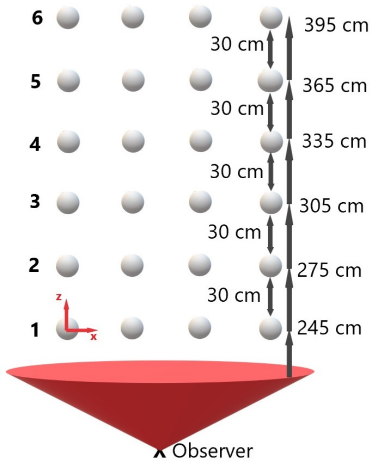


Figure 2: Schematic view of the proposed object configuration from the bird perspective (top view). The objects can be placed at 6 different, labelled depth levels with the constant distance difference between each level.

to explore the relationship between depth bias trends for each scene configuration type. The correlation is generally present, but is weak ($r = 0.199; p < .001$). Nevertheless, the trends of the higher depth saliency of the closer and farther objects are still obviously present, as studied by Desingh et al. [14] and Ramasamy et al. [61].

The first key to the internal validity of the comparative study is the consistency of the visual scene across different types of the environment. Therefore, two-dimensional and three-dimensional captures of the prepared scenes in the real environment were created to use them further for the two-dimensional widescreen projection and in virtual reality, respectively. The second key to the internal validity of each o the studies in a different type of the environment is the consistency of the control variables, namely background setup of the visually neutral scene, lighting conditions, position and direction of the participant's visual field across all participants and all visual scenes presented to them. We rely on the assumption that the internal validity was maintained due to the design of the methodology for the study, where the initial calibration of the participant's perception was triggered by freely looking at the environment with an internally valid initial calibration scene setup for the concrete type of the environment. We assume that the depth perception of the participant is then further dependent on this initial state and perception mechanisms estimate and guess the distance relative to the reference obtained during the calibration step [53].

Speaking about the external validity, we have to point out that the designed comparative study on the binocular depth perception under 3 different environmental conditions (reality, virtual reality, two-dimensional widescreen projection) was placed into laboratory conditions for each type of the environment. We propose the study in the laboratory conditions because of the motivation to contribute to the previous work of Laco et al. [37] and Wang et al. [70] by moving research in the field of quantification and modelling of the binocular depth perception a step further where control variables for each type of the environment can be maintained. Therefore, we claim that the results, findings and discussion are bound to the artificial scenes designed in the laboratory conditions where some depth cues helping the depth estimation and guessing could not play and, therefore, could not impact the attention tendencies of the participants. Moreover, we have to emphasize that we can speak about the limited external validity in the indoor environment with the artificial lighting and relatively small absolute distance of the farthest objects (395 centimetres by the design of the laboratory). Our work can be extended and realized in an unconstrained outdoor environment if the authors can bring up a solution for maintaining control variables unchanged to make the evaluation possible in the future.

6 Endogenous Bias of Emotional State on Visual Attention

There is an evidence that distractors and engagement factors, along with the motivation, have a clear impact on the visual attention [19, 55, 67, 59, 4, 56, 54]. Therefore, we decided to examine this open research topic further.

Emotions and their temporal state belong to the individual internal factors which act as distractors when it comes to the impact on visual attention, according to several studies[19, 55, 67, 59]. This fact is closely related to the broaden-and-build theory introduced by Fredrickson et al. [19]. This widely spread and accepted theory about the emotional bias of attention claims that positive emotions broaden attention significantly. It enlarges the set of visual stimuli the observer is able to be focused on and enhances the set of the possible reactions in a particular and individual way. Therefore, there is strong evidence of spreading attention when observing stimuli under emotionally non-neutral conditions. Moreover, individuality plays a significant role, and the correlation of the measured attention between examined subjects should be lower than under neutral conditions. Later on, Rowe et al.[64] proved the broaden-and-build theory claiming that a positive mood inducted more semantic associations related to the observed scene in comparison with the observers in a neutral mood. On the other hand, a few studies are not consistent with this theory when it comes to visual search tasks.

Internal motivation and engagement based on the emotional state of the observer is another individual factor with an impact on the attention [4, 56, 54]. There is a notice that emotion drives motivation and enhances visual performance[60, 5, 56]. Internal motivation can be a powerful aspect affecting attention, which suppresses other aspects, such as distractors, and causes even a lack of control over the attention[54]. Studies proved that fear is such a motivational aspect affecting the attention [4] and visual search. According to Ohman et al.[56], stimuli inducing fear were found by the observers much faster than neutral stimuli.

Surprisingly, engagement is an aspect with the opposite effect of motivation. Generally, lower engagement in the visual task is beneficial for the visual performance [65, 58, 32]. Smilek et al. [65] state that this fact may be based on the assumption that higher engagement suppresses the natural automatic visual perception process causing the visual system to perform worse and even causing the unwanted attentional blink, according to Olivers et al. and Jefferies et al.[58, 32].

Modelling the visual attention concerning the internal state of the observers (their emotional state, erosion rate, motivation and engagement rates) may be beneficial in various ways. Knowing this quantified information about the individual observers would help us suppress these factors' impact in the measured attention data while studying scene-driven aspects solely. Moreover, it could bring us completely new possibilities in the modelling of personalized visual attention [72, 73]. Some studies took into account the impact of emotions in the process of attention modelling [48, 15]. These proposals consider emotionally-tuned reactions to certain stimuli like snakes or flowers at the scene and train their models with the pictorial data labelled with emotionally-tuned responses on specific scene locations. However, there is a lack of models which would take into account the temporal internal state of the observer regardless of the pictorial stimuli. The impact of the temporal emotional state of the observers on their attention behaviour should be thoroughly examined to open up such possibilities.

We propose a methodology for the experimental eye-tracking study to obtain visual attention and visual performance data about a statistically relevant sample of subjects to be able to evaluate and discuss hypotheses $H2.1_0$ and $H2.2_0$. The experimental eye-tracking study should consist of these phases for each of the participating subjects:

1. the introductory instructions,
2. the emotion induction session with an expert in psychology,
3. the eye-tracking session.

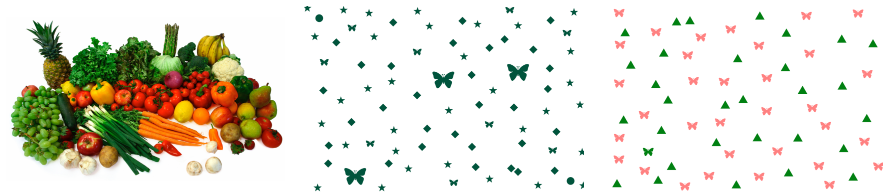
We used 40 emotionally-neutral images as the visual stimuli for the study. From these, 26 images were artificial images designed by the team of Laco et al. [1] (see Figure 4) and 14 were captions of natural scenes from the COCO dataset [47] (see Figure 3).

(a) (b)

Figure 3: Examples of the images from the COCO dataset [47] chosen for the free-viewing task.

The three types of visual search tasks designed for the eye-tracking study were:

1. **Search tasks:**
   (a) **FO-task:** *find* a specific *target object* among non-targets (e.g. find a triangle;



(a) FO task: find bananas. (b) FOA task: find circle. (c) FU taks: find unique object (green butterfly).

Figure 4: Examples of the images used for the visual search tasks [1].

14 tasks),

(b) **FOA-task:** *find any* one specific *target object* among non-targets where more objects meet the specific criteria (e.g. find any cloud; 5 tasks),

(c) **FU-task:** *find* a *unique object* whose visual attributes differ from other objects' attributes (9 tasks),

2. **V-task:** *free view* of an image (5 tasks),
3. **M-task:** *memorize* the image content (5 tasks).

A total of 39 participants were assigned to the emotionally neutral group (N-group), 43 participants were assigned to the emotionally affected group using positive memory recall (PI-group), and 24 participants were assigned to the emotionally affected group using music with positive affect (PM-group). Subjects belonging to the PI group ($M_{dn} = 2$) were significantly more positive than those in N group ($M_{dn} = 3$) ($U = 610, p = .024$) (lower M_{dn} means 'higher positivity and enthusiasm). Similarly, subjects from the PM group ($M_{dn} = 2,5$) were significantly more positive than those from N group ($M_{dn} = 3$) ($U = 312, p = .018$).

We analysed the statistical significance of the TCT metrics results over the N, PI and PM groups for each of the search task types separately using the one-way ANOVA test [10] and over the PI and PM groups with the control group separately for each of the search task types using the two-tailed t-test [10]. We observed the lowest statistical p-values only for the FO task between all of the groups ($F = 2.3865, p = 0.0977$), the FO task between the PM and N group ($t = 1.6099, p = 0.1133$) and the FU task between the PM and N group ($t = 1.8977, p = 0.0642$). Considerably better search performance of subjects was bound to the stimuli where the target differed from the non-target by a limited number of the bottom-up visual attributes. The statistics did not prove any significant difference between the examined groups in the means of the TCT distributions and, thus, we have to accept null hypothesis $H2.1_0$.

The lowest average fixation correlation was achieved by the subjects in the PM group for the FOA task types ($\bar{x} = 0.370; \sigma = 0.075$). Here, the broaden-and-build theory affected the subjects greatly as the distraction caused a wide broadening of the subjects' attention for a wider set of targets, mainly differing in many bottom-up attributes. Even though the fixation correlations are higher for the other task types, they are still low for accepting null hypothesis $H2.2_0$ for the M task type ($\bar{x} = 0.662; \sigma = 0.061$), FU task type ($\bar{x} = 0.628; \sigma = 0.077$) and V task ($\bar{x} = 0.513; \sigma = 0.169$). We conclude that the bottom-up attention of a subject under higher encouragement and distraction state, thus broadening the subject's attention, leads to different footprints than those of the non-affected subject. We consider null hypothesis $H2.2_0$ as rejected and, thus we accept alternative hypothesis $H2.2_1$.

7 Endogenous Bias of Individuality on Visual Attention

Personalized attention models have the potential to perform better than generalized attention models when predicting the attention tendencies of a certain subject [3, 23, 72, 42]. Along with one of the first publications related to personalized attention modelling, the extensive *Personalized Saliency Dataset* (PSD) was introduced by Xu et al. [72]. The dataset contains 1600 images of natural scenes from various sources. Each image was observed for 3 seconds by a sample of 30 participants during the extensive eye-tracking sessions.

It has been already demonstrated that deep learning techniques are beneficial to use for modelling personalized attention [74, 46]. Yu et al. [74] in their work aimed to generate personalized saliency maps using Deep Convolutional Adversarial Network, which takes as an input only the image data and corresponding saliency maps of the individuals. They conclude that the proposed network efficiently generates saliency maps with a focus on personalization based on age and the natively spoken language of the subjects. Similarly, Lin et al. [46] proposed a personalized model based on deep learning and a convolutional network with shared feature extraction layers and two input streams consisting of a generalized saliency map for the given image and the preference fitting stream as the representation of the individual's characteristics. There is a conclusion stating that the more biased the individual's attention compared to the generalized model, the better the performance of their personalized model is. Therefore, it is reasonable to put great effort into personalized attention modelling and its applications during the following years.

We build our proposal for attention model personalization on the assumption that standard state-of-the art attention models are able to capture attention tendencies for the major population with high precision [7]. An overview of the personalization methodology for generating a personalized attention model based on a generalized visual attention model for a specific subject can be found in Figure 5.

We decided to prove our methodology proposal and assess our hypothesis over the MSI-NET architecture [36] scoring one of the highest scores in evaluation metrics in the MIT/Tuebingen saliency benchmark. The saliency metrics we propose to evaluate are according to the recommendation of Borji et al. [7], and Bylinskii et al. [13]: AUC-Judd, AUC-Borji, NSS, SIM, CC, KL-Div. For the discussion purposes, we statistically generalized the evaluation data in the form of an average of the improvements across the personalized attention models using these saliency metrics. The aggregated results are visualized in Table 1.

We claim that quantitative results, their distribution and qualitative analysis undoubtedly indicate that the proposed methodology for personalisation of the attention models is beneficial for more precise predictions for the statistical majority of the relevant 10-

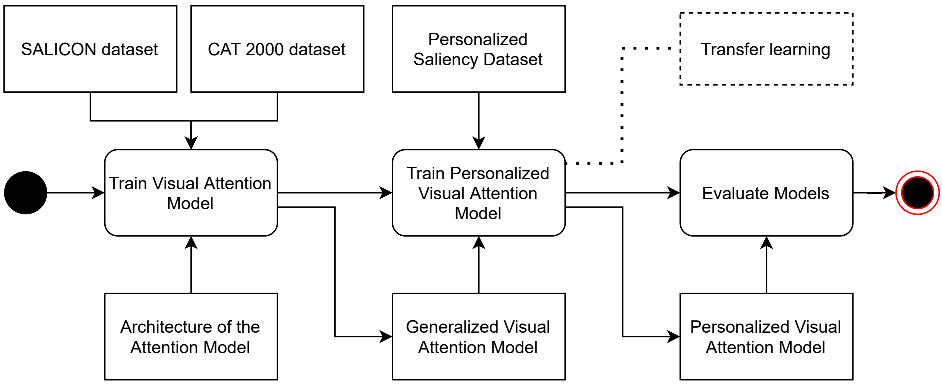


Figure 5: Visualization of the general claim about creating visual attention models where the nature of the data defines the resulting attention model with a shared architecture able to extract visual attention tendencies.

fold cross-validated subset of subjects in the evaluation dataset with certain limitations discussed in this paragraph. Therefore, we claim that the discussed results reject null hypothesis $H3.1_0$ and we consider alternative hypothesis $H3.1_1$ as accepted.

On the other hand, the Shapiro-Wilk test rejected the hypothesis $H3.2_0$ ($\alpha = 0.05; p < .01$) that the AUC-Judd distribution of the improvements of the visual attention predictions generated by the personalized visual attention model over the generalized visual attention model is from the normal distribution (see Figure 6). Therefore, we should reject null hypothesis $H3.2_0$ and accept the alternative hypothesis $H3.2_1$.

Based on the quantitative evaluation, we prepared representative visualizations of the

Metric	Unisal [16]	Generalized MSI-Net [36]	Personalized MSI-Net
AUC Judd(↑)	0.8847	0.8906	**0.8978**
AUC Borji(↑)	0.8251	**0.8399**	0.8369
NSS (↑)	2.2923	2.2667	**2.4605**
SIM (↑)	0.5255	0.5291	**0.5624**
CC (↑)	0.6369	0.6305	**0.6692**
KLD (↓)	0.9730	0.8709	**0.7759**
IG GSM(↑)	-0.1934	0.0000	**0.1364**

Table 1: Metric results of the proposed personalized model predictions averaged throughout participants compared with the basal model on the PSD dataset [72].

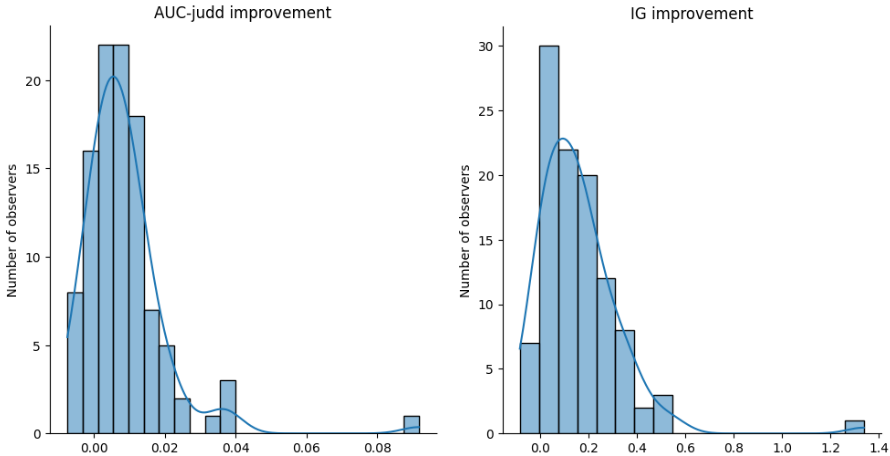


Figure 6: Distributions of the AUC-Judd and IG GSM metric improvements of the visual attention predictions generated by the personalized visual attention model over the generalized visual attention model for the PSD dataset [72].

qualitative results to discuss the performance of both versions of the personalized models for specific subjects. We can see obvious benefits of the model personalization in Figure 7 which was chosen as the positive extreme in the improvement of the personalized model over the general one. On the other hand, the opposite extreme is visualized in Figure 8.

From the visualizations in Figures 7 and 7, we can observe that the personalization improves the model greatly in the mean of reducing false negatives in predictions over the ground-truth data. This is also reflected by the improvements in the IG GSM metric, which is sensitive for such a model behaviour. It comes out from the application praxis where omitting some important regions from images (false negatives) is worse than marking more image regions as important (false positives). However, the consequence of the improvement in reducing false negatives is obviously in the introduction of more false positives than the generalized model predictions suffered from. This balance is then penalized by some of the common metrics such as AUC-Borji, or NSS. Thus, our personalization approach resulted in a worse average AUC-Borji score. However, by qualitative analysis of the mentioned Figures, we claim that the loss on this metric has no negative impact on the general performance of the personalized models.

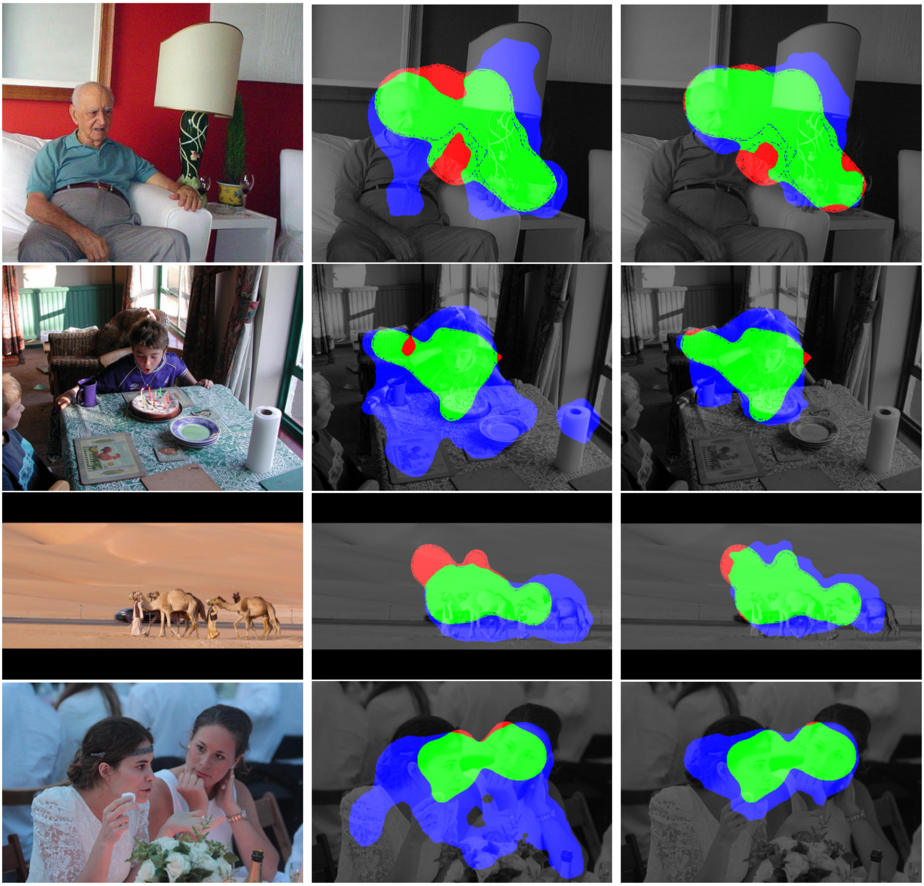


Figure 7: Examples of the qualitative improvements of the personalized saliency predictions using our proposed personalization methodology (on the **right** side) in comparison with the generalized prediction of the basal generalized visual attention model (in the **middle**) over the input visual stimuli (on the **left** side) from the PSD dataset [72]. The green colour is mapped to true positives, the red colour to false negatives and the blue colour to false positives.

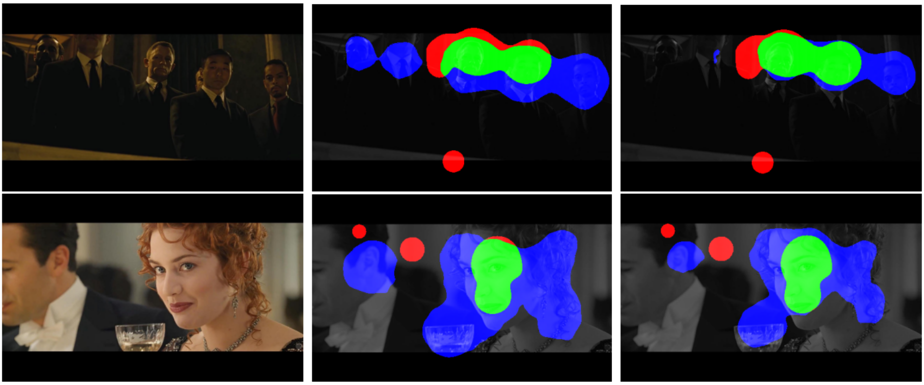


Figure 8: Examples of the limitations of the personalized saliency predictions using our proposed personalization methodology (on the **right** side) in comparison with the generalized prediction of the basal non-personalized model (in the **middle**) over the input visual stimuli (on the **left** side) from the PSD dataset [72]. The green colour is mapped to true positives, the red colour to false negatives and the blue colour to false positives. We stress out mainly the model's confusion for certain types of image inputs where the prediction is very spread towards the background and the limitations towards understanding higher top-down context present in the image.

8 Conclusions

We analysed principles of visual perception and attention to kick off our research in the growing application field of visual attention modelling. We recently identified open research topics in the field and novel perspective research directions in visual attention modelling. Based on the identified topics within the field of visual attention modelling, we formulated three research theses matching our research goals in the form of multiple hypotheses. We initiated thorough research of the state-of-the-art directly related to the research topics regarding the identified theses. For each of the theses, we identified the actual motivation for solving the problem in the application fields in the future, formulated our assumptions related to the identified hypotheses, proposed a research methodology to solve and thoroughly discuss and assess the hypotheses using advanced principles of computer vision, statistics for the experimental studies, and deep-learning techniques. We proposed the evaluation methodology for each of the hypotheses. The results obtained using the proposed methodologies were presented and discussed based on the evaluation methodology. The hypotheses for each thesis were assessed and evaluated. Finally, we discussed applying our research results to the state-of-the-art in the field and discussed the possible future applications. In conclusion, this thesis is a complete research work on the three chosen research topics, which push the knowledge in the field towards further progress.

In our future work, we aim to continue further the research related to individuality in attention modelling. We aim to start a cooperation with experts in the field of human-computer interaction and medicine diagnostics to discuss the further necessary steps to take to apply our research into praxis.

Bibliography

2. ACHANTA, Radhakrishna; SÜSSTRUNK, Sabine. Saliency detection for content-aware image resizing. In: *Image Processing (ICIP), 2009 16th IEEE International Conference on*. IEEE, 2009, pp. 1005–1008.
3. ANDERSON, Nicola C; ANDERSON, Fraser; KINGSTONE, Alan; BISCHOF, Walter F. A comparison of scanpath comparison methods. *Behavior research methods*. 2015, vol. 47, no. 4, pp. 1377–1392.
4. ARMONY, Jorge L; DOLAN, Raymond J. Modulation of spatial attention by fear-conditioned stimuli: an event-related fMRI study. *Neuropsychologia*. 2002, vol. 40, no. 7, pp. 817–826.
5. BARBOT, Antoine; CARRASCO, Marisa. Emotion and anxiety potentiate the way attention alters visual appearance. *Scientific reports*. 2018, vol. 8, no. 1, p. 5938.
6. BECKER, Mark W; LEINENGER, Mallorie. Attentional selection is biased toward mood-congruent stimuli. *Emotion*. 2011, vol. 11, no. 5, p. 1248.
7. BORJI, Ali; CHENG, Ming-Ming; HOU, Qibin; JIANG, Huaizu; LI, Jia. Salient object detection: A survey. *Computational visual media*. 2019, vol. 5, no. 2, pp. 117–150.
8. BORJI, Ali; ITTI, Laurent. State-of-the-art in visual attention modeling. *IEEE transactions on pattern analysis and machine intelligence*. 2013, vol. 35, no. 1, pp. 185–207.
9. BORJI, Ali; ITTI, Laurent. Cat2000: A large scale fixation dataset for boosting saliency research. *arXiv preprint arXiv:1505.03581*. 2015.
10. BOX, G; G. HUNTER, William; STUART HUNTER, J. *Statistics For Experimenters*. Vol. 21. 1978. Available from DOI: `10.2307/1267766`.
11. BRUCE, Neil; TSOTSOS, John. Saliency based on information maximization. In: *Advances in neural information processing systems*. 2006, pp. 155–162.
12. BRUCE, Neil DB; TSOTSOS, John K. Saliency, attention, and visual search: An information theoretic approach. *Journal of vision*. 2009, vol. 9, no. 3, pp. 5–5.
13. BYLINSKII, Zoya; JUDD, Tilke; OLIVA, Aude; TORRALBA, Antonio; DURAND, Frédo. What do different evaluation metrics tell us about saliency models? *IEEE transactions on pattern analysis and machine intelligence*. 2018, vol. 41, no. 3, pp. 740–757.
14. DESINGH, Karthik; KRISHNA K, Madhava; RAJAN, Deepu; JAWAHAR, CV. Depth really Matters: Improving Visual Salient Region Detection with Depth. In: 2013, pp. 98.1–98.11. ISBN 1-901725-49-9. Available from DOI: `10.5244/C.27.98`.

15. DING, Xinmiao; HUANG, Lulu; LI, Bing; LANG, Congyan; HUA, Zhen; WANG, Yuling. A Novel Emotional Saliency Map to Model Emotional Attention Mechanism. In: *International Conference on Multimedia Modeling*. Springer, 2016, pp. 197–206.

16. DROSTE, Richard; JIAO, Jianbo; NOBLE, J Alison. Unified image and video saliency modeling. In: *European Conference on Computer Vision*. Springer, 2020, pp. 419–435.

17. DUCHOWSKI, Andrew T. Eye tracking methodology. *Theory and practice*. 2007, vol. 328, no. 614, pp. 2–3.

18. FINLAYSON, Nonie J.; REMINGTON, Roger W.; RETELL, James D.; GROVE, Philip M. Segmentation by depth does not always facilitate visual search. *Journal of Vision*. 2013, vol. 13, no. 8, pp. 11–11. ISSN 1534-7362. Available from DOI: 10.1167/13.8.11.

19. FREDRICKSON, Barbara L. The broaden-and-build theory of positive emotions. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2004, vol. 359, no. 1449, p. 1367.

20. GAO, Dashan; MAHADEVAN, Vijay; VASCONCELOS, Nuno. On the plausibility of the discriminant center-surround hypothesis for visual saliency. *Journal of vision*. 2008, vol. 8, no. 7, pp. 13–13.

21. GOLDSTEIN, E Bruce. *Encyclopedia of perception*. Vol. 1. Sage, 2010.

22. GOLDSTEIN, E Bruce; BROCKMOLE, James. *Sensation and perception*. 8th ed. Cengage Learning, 2016. ISBN 0495760501.

23. GYGLI, Michael; GRABNER, Helmut; RIEMENSCHNEIDER, Hayko; NATER, Fabian; VAN GOOL, Luc. The interestingness of images. In: *Proceedings of the IEEE International Conference on Computer Vision*. 2013, pp. 1633–1640.

24. HAREL, Jonathan; KOCH, Christof; PERONA, Pietro. Graph-based visual saliency. In: *Advances in neural information processing systems*. 2007, pp. 545–552.

25. HOU, Qibin; CHENG, Ming-Ming; HU, Xiaowei; BORJI, Ali; TU, Zhuowen; TORR, Philip HS. Deeply supervised salient object detection with short connections. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 3203–3212.

26. HOU, Xiaodi; ZHANG, Liqing. Saliency detection: A spectral residual approach. In: *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–8.

27. ITTI, Laurent; BALDI, Pierre. A principled approach to detecting surprising events in video. In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. IEEE, 2005, vol. 1, pp. 631–637.

28. ITTI, Laurent; DHAVALE, Nitin; PIGHIN, Frederic. Realistic avatar eye and head animation using a neurobiological model of visual attention. In: *Applications and Science of Neural Networks, Fuzzy Systems, and Evolutionary Computation VI*. International Society for Optics and Photonics, 2003, vol. 5200, pp. 64–79.

29. ITTI, Laurent; KOCH, Christof. A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision research*. 2000, vol. 40, no. 10-12, pp. 1489–1506. Available from DOI: `DOI:10.1016/S0042-6989(99)00163-7`.

30. ITTI, Laurent; KOCH, Christof; NIEBUR, Ernst. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*. 1998, no. 11, pp. 1254–1259. Available from DOI: `10.1109/34.730558`.

31. JANSEN, Lina; ONAT, Selim; KÖNIG, Peter. Influence of disparity on fixation and saccades in free viewing of natural scenes. *Journal of Vision*. 2009, vol. 9, no. 1, pp. 29–29. Available from DOI: `10.1167/9.1.29`.

32. JEFFERIES, Lisa N; SMILEK, Daniel; EICH, Eric; ENNS, James T. Emotional valence and arousal interact in attentional control. *Psychological Science*. 2008, vol. 19, no. 3, pp. 290–295.

33. JIANG, Ming; HUANG, Shengsheng; DUAN, Juanyong; ZHAO, Qi. Salicon: Saliency in context. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1072–1080.

34. KOMOGORTSEV, Oleg V.; GOBERT, Denise V.; JAYARATHNA, Sampath; KOH, Do Hyong; GOWDA, Sandeep M. Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors. *IEEE Transactions on Biomedical Engineering*. 2010, vol. 57, no. 11, pp. 2635–2645. Available from DOI: `10.1109/TBME.2010.2057429`.

35. KRISHNA, Onkar; HELO, Andrea; RÄMÄ, Pia; AIZAWA, Kiyoharu. Gaze distribution analysis and saliency prediction across age groups. *PloS one*. 2018, vol. 13, no. 2, e0193149.

36. KRONER, Alexander; SENDEN, Mario; DRIESSENS, Kurt; GOEBEL, Rainer. Contextual encoder–decoder network for visual saliency prediction. *Neural Networks*. 2020, vol. 129, pp. 261–270.

37. LACO, Miroslav; BENESOVA, Wanda. Depth in the visual attention modelling from the egocentric perspective of view. In: VERIKAS, Antanas; NIKOLAEV, Dmitry P.; RADEVA, Petia; ZHOU, Jianhong (eds.). *Eleventh International Conference on Machine Vision (ICMV 2018)*. SPIE, International Society for Optics and Photonics, 2019, vol. 11041, pp. 329–339. Available from DOI: `10.1117/12.2523059`.

1. LACO, Miroslav; POLATSEK, Patrik; DEKRÉT, Šimon; BENESOVA, Wanda; BARÁNKOVÁ, Martina; STRNÁDELOVÁ, Bronislava; KORÓNIOVÁ, Jana; GABLÍKOVÁ, Mária. Effects of individual's emotions on saliency and visual search. *The Visual Computer*. 2020. ISSN 1432-2315. Available from DOI: `10.1007/s00371-020-01912-7`.

38. LACO, Miroslav; POLATSEK, Patrik; POLLÁKOVÁ, Jana; KUL'BAK, Daniel; KAPEC, Peter; BENESOVA, Wanda. Comparative Study on the Egocentric Depth Perception under Various Environment Conditions. *Submitted to: Special Issue on Egocentric Perception in the journal 'IEEE Transactions on Pattern Analysis and Machine Intelligence'. Waits for re-submission.* 2020.

39. LAND, Michael; TATLER, Benjamin. *Looking and acting: vision and eye movements in natural behaviour*. Oxford University Press, 2009.

40. LANG, Congyan; NGUYEN, Tam V; KATTI, Harish; YADATI, Karthik; KANKANHALLI, Mohan; YAN, Shuicheng. Depth matters: Influence of depth cues on visual saliency. In: *European conference on computer vision*. Springer, 2012, pp. 101–115. Available from DOI: `10.1007/978-3-642-33709-3_8`.

41. LEE, Gayoung; TAI, Yu-Wing; KIM, Junmo. Deep saliency with encoded low level distance map and high level features. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 660–668.

42. LI, Aoqi; CHEN, Zhenzhong. Personalized visual saliency: Individuality affects image perception. *IEEE Access*. 2018, vol. 6, pp. 16099–16109.

43. LI, Guanbin; XIE, Yuan; LIN, Liang; YU, Yizhou. Instance-level salient object segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 2386–2395.

44. LI, Guanbin; YU, Yizhou. Deep contrast learning for salient object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 478–487.

45. LI, Hongyang; CHEN, Jiang; LU, Huchuan; CHI, Zhizhen. CNN for saliency detection with low-level feature integration. *Neurocomputing*. 2017, vol. 226, pp. 212–220.

46. LIN, Sikun; HUI, Pan. Where's YOUR focus: Personalized Attention. *arXiv preprint arXiv:1802.07931*. 2018.

47. LIN, Tsung-Yi; MAIRE, Michael; BELONGIE, Serge; HAYS, James; PERONA, Pietro; RAMANAN, Deva; DOLLÁR, Piotr; ZITNICK, C Lawrence. Microsoft coco: Common objects in context. In: *European conference on computer vision*. Springer, 2014, pp. 740–755.

48. LIU, Huiying; XU, Min; WANG, Jinqiao; RAO, Tianrong; BURNETT, Ian. Improving visual saliency computing with emotion intensity. *IEEE transactions on neural networks and learning systems*. 2016, vol. 27, no. 6, pp. 1201–1213.

49. LIU, Tie; SUN, Jian; ZHENG, Nan-Ning; TANG, Xiaoou; SHUM, Heung-Yeung. Learning to detect a salient object. In: *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–8.

50. MAHADEVAN, Vijay; VASCONCELOS, Nuno. Spatiotemporal saliency in dynamic scenes. *IEEE transactions on pattern analysis and machine intelligence*. 2010, vol. 32, no. 1, pp. 171–177.

51. MAKI, Atsuto; NORDLUND, Peter; EKLUNDH, Jan-Olof. Attentional scene segmentation: integrating depth and motion. *Computer Vision and Image Understanding*. 2000, vol. 78, no. 3, pp. 351–373. Available from DOI: `10.1006/cviu.2000.0840`.

52. MANCAS, Matei et al. *Computational attention towards attentive computers*. Presses univ. de Louvain, 2007.

53. MATHER, George. Image blur as a pictorial depth cue. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 1996, vol. 263, no. 1367, pp. 169–172. Available from DOI: `10.1098/rspb.1996.0027`.

54. MIYAZAWA, Shiho; IWASAKI, Syoichi. Effect of negative emotion on visual attention: Automatic capture by fear-related stimuli. *Japanese Psychological Research*. 2009, vol. 51, no. 1, pp. 13–23.

55. MOST, Steven B; SMITH, Stephen D; COOTER, Amy B; LEVY, Bethany N; ZALD, David H. The naked truth: Positive, arousing distractors impair rapid target perception. *Cognition and emotion*. 2007, vol. 21, no. 5, pp. 964–981.

56. ÖHMAN, Arne; FLYKT, Anders; ESTEVES, Francisco. Emotion drives attention: detecting the snake in the grass. *Journal of experimental psychology: general*. 2001, vol. 130, no. 3, p. 466.

57. OLESOVA, Veronika; BENESOVA, Wanda; POLATSEK, Patrik. Visual attention in egocentric field-of-view using RGB-D data. In: *Ninth International Conference on Machine Vision (ICMV 2016)*. International Society for Optics and Photonics, 2017, vol. 10341, 103410T. Available from DOI: `10.1117/12.2268617`.

58. OLIVERS, Christian NL; NIEUWENHUIS, Sander. The beneficial effect of concurrent task-irrelevant mental activity on temporal attention. *Psychological science*. 2005, vol. 16, no. 4, pp. 265–269.

59. PÊCHER, Christelle; LEMERCIER, Céline; CELLIER, Jean-Marie. Emotions drive attention: Effects on driver's behaviour. *Safety Science*. 2009, vol. 47, no. 9, pp. 1254–1259.

60. PHELPS, Elizabeth A; LING, Sam; CARRASCO, Marisa. Emotion facilitates perception and potentiates the perceptual benefits of attention. *Psychological science*. 2006, vol. 17, no. 4, pp. 292–299.

61. RAMASAMY, Celambarasan; HOUSE, Donald H; DUCHOWSKI, Andrew T; DAUGHERTY, Brian. Using eye tracking to analyze stereoscopic filmmaking. In: *SIGGRAPH'09: Posters*. ACM, 2009, p. 28. Available from DOI: 10.1145/1599301.1599329.

62. RAYNER, Keith. Eye movements and attention in reading, scene perception, and visual search. *The quarterly journal of experimental psychology*. 2009, vol. 62, no. 8, pp. 1457–1506.

63. ROBERTS, Katherine L; ALLEN, Harriet A; DENT, Kevin; HUMPHREYS, Glyn W. Visual search in depth: The neural correlates of segmenting a display into relevant and irrelevant three-dimensional regions. *NeuroImage*. 2015, vol. 122, pp. 298–305. Available from DOI: 10.1016/j.neuroimage.2015.07.052.

64. ROWE, Gillian; HIRSH, Jacob B; ANDERSON, Adam K. Positive affect increases the breadth of attentional selection. *Proceedings of the National Academy of Sciences*. 2007, vol. 104, no. 1, pp. 383–388.

65. SMILEK, Daniel; ENNS, James T; EASTWOOD, John D; MERIKLE, Philip M. Relax! Cognitive strategy influences visual search. *Visual Cognition*. 2006, vol. 14, no. 4-8, pp. 543–564.

66. SQUIRE, L.R.; BLOOM, F.E.; SPITZER, N.C.; GAGE, F.; ALBRIGHT, T. *Encyclopedia of Neuroscience*. Elsevier Science, 2009. No. 1. ISBN 9780080963938.

67. TALARICO, Jennifer M; BERNTSEN, Dorthe; RUBIN, David C. Positive emotions enhance recall of peripheral details. *Cognition and Emotion*. 2009, vol. 23, no. 2, pp. 380–398.

68. TORRALBA, Antonio; OLIVA, Aude; CASTELHANO, Monica S; HENDERSON, John M. Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychological review*. 2006, vol. 113, no. 4, p. 766.

69. TREISMAN, Anne M; GELADE, Garry. A feature-integration theory of attention. *Cognitive psychology*. 1980, vol. 12, no. 1, pp. 97–136.

70. WANG, Junle; LE CALLET, Patrick; TOURANCHEAU, Sylvain; RICORDEL, Vincent; DA SILVA, Matthieu Perreira. Study of depth bias of observers in free viewing of still stereoscopic synthetic stimuli. *Journal of Eye Movement Research*. 2012, vol. 5, no. 5, pp–1.

71. WANG, Lijun; LU, Huchuan; RUAN, Xiang; YANG, Ming-Hsuan. Deep networks for saliency detection via local estimation and global search. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, pp. 3183–3192.

72. XU, Yanyu; GAO, Shenghua; WU, Junru; LI, Nianyi; YU, Jingyi. Personalized saliency and its prediction. *IEEE transactions on pattern analysis and machine intelligence*. 2018, vol. 41, no. 12, pp. 2975–2989.

73. XU, Yanyu; LI, Nianyi; WU, Junru; YU, Jingyi; GAO, Shenghua. Beyond Universal Saliency: Personalized Saliency Prediction with Multi-task CNN. In: *IJCAI*. 2017, pp. 3887–3893.

74. YU, Bingqing; CLARK, James J. Personalization of saliency estimation. *arXiv preprint arXiv:1711.08000*. 2017.

75. ZHANG, Lingyun; TONG, Matthew H; MARKS, Tim K; SHAN, Honghao; COTTRELL, Garrison W. SUN: A Bayesian framework for saliency using natural statistics. *Journal of vision*. 2008, vol. 8, no. 7, pp. 32–32.

76. ZHAO, Rui; OUYANG, Wanli; LI, Hongsheng; WANG, Xiaogang. Saliency detection by multi-context deep learning. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015, pp. 1265–1274.

A Publications of the Author

Zoznam publikačnej činnosti
Autor: Laco, Miroslav

ADC Vedecké práce v zahraničných karentovaných časopisoch

ADC01 LACO, Miroslav - POLATSEK, Patrik - DEKRÉT, Šimon - BENEŠOVÁ, Vanda - BARÁNKOVÁ, Martina - STRNADELOVÁ, Bronislava - KORONIOVÁ, Jana - GÁBLIKOVÁ, Mária. Effects of individual's emotions on saliency and visual search. In *The Visual Computer*. Vol. 37, no. 6 (2021), s. 1581-1592. ISSN 0178-2789 (2020: 2.601 - IF, Q2 - JCR Best Q, 0.316 - SJR, Q3 - SJR Best Q). V databáze: WOS: 000553760200003 ; SCOPUS: 2-s2.0-85088793866 ; DOI: 10.1007/s00371-020-01912-7.

AFC Publikované príspevky na zahraničných vedeckých konferenciách

AFC01 IVANOVÁ, Lenka - LACO, Miroslav - BENEŠOVÁ, Vanda. Unsupervised clustering-based analysis of the measured eye-tracking data. In *Fourteenth International Conference on Machine Vision (ICMV 2021)*. 1. vyd. Washington : SPIE - The International Society for Optical Engineering, 2022, S. [1-8], art. no. 1208408. ISSN 0277-786X. ISBN 9781510650442.

AFC02 LACO, Miroslav - BENEŠOVÁ, Vanda. Depth in the visual attention modelling from the egocentric perspective of view. In *Proceedings SPIE 11041, Eleventh International Conference on Machine Vision (ICMV 2018), 2018, Munich, Germany*. 1. vyd. Bellingham : SPIE - The International Society for Optical Engineering, 2019, S. ISBN 9781510627482. V databáze: WOS: 000468216400043 ; SCOPUS: 2-s2.0-85063439711.

AFC03 LACO, Miroslav - POLATSEK, Patrik - BENEŠOVÁ, Vanda. Depth Perception Tendencies on a Widescreen Display: An Experimental Study. In *Proceedings SPIE 11433, Twelfth International Conference on Machine Vision (ICMV 2019), 2019, Amsterdam, The Netherlands*. 1. vyd. Bellingham : SPIE - The International Society for Optical Engineering, 2020, S. 1-8. ISSN 0277-786X (print). ISBN 9781510636439 (print). V databáze: SCOPUS: 2-s2.0-85081179840 ; WOS: 000540890900061.

AFC04 POLLÁKOVÁ, Jana - LACO, Miroslav - BENEŠOVÁ, Vanda. Depth Perception Tendencies in the 3-D Environment of Virtual Reality. In *Computer Vision and Graphics : International Conference, ICCVG 2020, Warsaw, Poland, September 14–16, 2020, Proceedings*. 1. ed. Cham : Springer Nature, 2020, S. 142-150. ISSN 0302-9743 (print). ISBN 978-3-030-59005-5 (print). V databáze: SCOPUS: 2-s2.0-85091338527 ; DOI: 10.1007/978-3-030-59006-2_13.

Štatistika: kategória publikačnej činnosti

ADC	Vedecké práce v zahraničných karentovaných časopisoch	1
AFC	Publikované príspevky na zahraničných vedeckých konferenciách	4
Súčet		**5**