

iMap: A Novel Method
for Statistical Fixation Mapping of Eye Movement data

Roberto Caldara^{1*,2} and Sébastien Miellet²

¹ Department of Psychology, University of Fribourg, Switzerland

² Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, University of Glasgow, United Kingdom

Correspondence should be addressed to:

Roberto Caldara

Department of Psychology

University of Fribourg

Rue de Faucigny 2

1700 Fribourg

Switzerland

+41 (0)26 300 76 68

roberto.caldara@unifr.ch

Short title: *iMap*

Abstract

Eye movement data analyses are commonly based on the probability of occurrence of saccades and fixations (and their characteristics) in given Regions of Interest (ROIs). In this paper, we introduce an alternative method – *iMap* – to compute statistical fixation maps of eye movements based on an approach inspired by methods used in functional Magnetic Resonance Imaging. Importantly, *iMap* does not require the *a priori* segmentation of the experimental images into ROIs. With *iMap*, fixation data are first smoothed by convoluting Gaussian kernels to generate 3D fixation maps. This procedure embodies the eye-tracker accuracy but can also be used, by setting the Gaussian kernel, to represent acuity or attentional constraints. In addition, the smoothed fixation data generated by *iMap* conform to the assumptions of the robust statistical Random Field Theory (RFT) approach applied thereafter to assess significant fixation spots and differences across 3D fixation maps. The RFT corrects for the multiple statistical comparisons generated by the numerous pixels constituting the digital images. To illustrate the processing steps of *iMap*, we provide real eye movement data on face, visual scene and memory processing. The *iMap* Matlab toolbox is editable and freely available to download online.

The human visual system is equipped with the most sophisticated machinery to effectively adapt to the visual world. *Where, when* and *how* human eyes are moved to gather information to adapt to the visual environment has been a question that has fascinated scientists for more than a century. Javal (1879) coined the term saccade to describe the rapid movement of the eyes produced during reading, an oculomotor phenomenon identified by Hering (1879) and Lamare (1892) during this period. However, a comprehensive sense of the very nature of those ballistic movements, the description of the use of fixations to gather the relevant information to solve the task at hand and the scientific definition of saccades, came with Dodge (1916) and the development of photographic techniques for recording corneal reflections. This novel recording approach paved the way to the scientific study of eye movements (see Wade, Tatler, & Heller, 2003).

Guy T. Buswell (1935) published the first systematic study on *How People Look at Pictures: A Study of The Psychology of Perception in Art*. Buswell observed that trained and untrained artists deployed similar fixation patterns to analyze paintings. All observers shared a similar oculomotor behaviour, deploying initial short fixations over the main features of the paintings, which were subsequently followed by a series of longer fixations. Interestingly, when fixations were collapsed across observers, they highlighted areas containing salient or diagnostic parts of the images. Critically, these observations revealed that eye movements are not randomly sampling the visual input space, but are effective to solve problems in visual cognition. This work was then followed by a series of studies, leading in particular to the seminal work of Yarbus (1965), which extensively showed similar findings for diverse visual objects. Yarbus showed more importantly how top-down factors modulate the eye movement strategy deployed by observers to gather information from the very same picture, attracting attention in the wider scientific community¹. Since then, we have witnessed an explosion of

¹ Note that Yarbus recorded only a single observer for demonstrating task effects on eye movement patterns.

eye movement studies in language and visual scenes (see Rayner, 1998; Rayner, 2009), as well as studies with clinical populations (see Van Gompel, Fischer, Murray, & Hill, 2007).

Nowadays, the development of new technologies that have increased the precision, the ease and affordability of eye tracking devices, has significantly impacted on the prominence of eye movement research. There is also an increasing awareness in the scientific community of the need to control eye movements during any experiment in vision (e.g., Yuval-Greenberg, Tomer, Keren, Nelken, & Deouell, 2008). These observations lead to the prediction that eye movement research will continue to gain importance in the future within the vision sciences community.

Scientific disciplines rely on their own specific metrics. Eye movement studies generate a quantity of rich data, which traditionally have largely relied on measures from two events characterizing *where*, *when* and *how* the eyes gather information from the visual world: *saccades* (i.e., their latency, amplitude, direction and occurrence over the time course) and *fixations* (i.e., location and duration), as well as many measures derived from these events such as cumulative saccade length, pupil dilation, etc.. To characterize and isolate statistical differences in the eye movements deployed to process visual inputs, the large majority of the eye movement literature has used a Region or Area of Interests (ROI or AOI) approach. Strictly and formally speaking, every single pixel of a digital image could be considered as a variable of interest to measure the occurrence of *saccades* and *fixations*, which results in a complex multidimensional space. The goal of segmentation is to reduce the visual input space (usually defined by thousands or millions of pixels) of the digital images used during the experimental tasks into something that is meaningful and easier to analyze. Image segmentation is usually based on a mixture of low-level boundaries of the object/feature shapes (lines, curves, etc.) present in the digital images, and high-level, semantic *a priori* that experimenters have about the parts constituting a particular visual

object. These boundaries are typically related to object/feature locations. More precisely, image segmentation is the process during which pixels are clustered together and assigned a label, such that pixels sharing a similar label would also share particular visual characteristics or semantic properties. This process results in a set of regions that collectively cover the entire image. Once the images used in a particular experiment are segmented into ROIs, descriptive eye movements are then measured, with the probability of fixation/saccade and their respective characteristics (i.e., number, duration, amplitude, etc.) calculated for each of the defined ROIs. The measures obtained for those metrics are thereafter submitted to conventional statistical analyses.

So far it has often been implicitly assumed that the ROIs are optimally representing the visual categories present in the visual input space (i.e., for face processing: the eye, the nose and mouth regions). However, segmenting visual inputs into ROIs is constrained by *subjective* evaluations, which is – by definition – problematic from a scientific point of view. For instance, how should the borders of a visual region representing the human eyes be concretely defined? Should both human eyes be considered as a single region? Should the pixels outside the eye sclera be considered as belonging to the eye region or not? If this is the case, how many pixels of the skin should be included? Should the pixels of this region be included by using a curvilinear, elliptical or rectangular shape? Obviously, beside an *objective* definition of a ROI based on the human sclera, the remaining options used to define ROIs for an eye region do not have an *objective* answer. For this reason, there is great variability between eye movement studies in the definition of ROIs representing the very same information: for instance, the eye region in faces (e.g., Barton, Radcliffe, Cherkasova, Edelman, & Intriligator, 2006; Henderson, Williams, & Falk, 2005; Orban de Xivry, Ramon, Lefevre, & Rossion, 2008). The same difficulties generalize to all the visual inputs. For instance, in the domain of visual scenes, it is difficult to define the objects to be included in

the ROI. Is it appropriate to determine different regions for a human body embedded in a visual scene (i.e., head, neck, hands, etc.), or would a unique shape defining the body be more appropriate? Should a ROI be strictly defined by using the edges of an object? The physical boundaries of objects are usually used to define ROIs. This choice appears sensible, however, under certain circumstances using ROIs might not be appropriate to thoroughly and effectively capture the eye movement behaviour (see Figure 1).

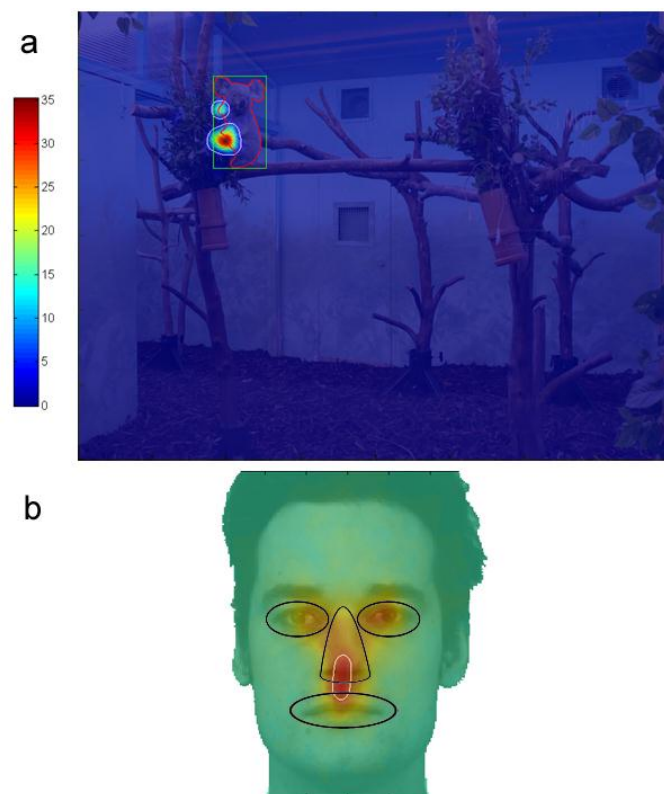


Figure 1. *Panel a* shows an example extracted from an item of an animal visual search task used in Miellet et al.'s study (2010). The red contour represents a ROI based on the edges of the panda, a rule routinely used in the eye movement literature relying on ROIs. The white contour shows areas of the visual scene that resulted in being significantly fixated above chance level with *iMap*. The centre of gravity of the location driving the majority of fixations in this image is located outside the ROI. This perceptual bias used by the observers, would be therefore inaccurately reported by probability of fixation analyses based on ROIs. *Panel b* illustrates the difficulty in defining *a priori* ROIs in face processing studies (data from Caldara et al., 2010). The white contour shows significant area according to *iMap*, blue contours show example of ROI as commonly used in the literature. In this example, it is difficult to *objectively* attribute the intermediary fixations to the nose or mouth regions.

But the more critical consequences of this drawback rely on the fact that the *subjective* criteria used to define ROIs compromise the potential to replicate findings across studies. Since ROIs are based on qualitative and quantitative *subjective* evaluations of the experimenters, they engender *natural* variations across authors, which in some cases thus lead to difficulty in generalizing observations across studies. Note that there are other potential problems of using ROIs has been also discussed in neuroimaging, particularly circular analyses and “double dipping” - the use of the same data set for selection and selective analysis (see Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). Therefore, this factor alone could potentially explain the absence of consistent effects across studies reported in the eye movement literature and points towards a methodological problem.

To overcome these limitations, we introduce a novel robust data-driven technique that does not require an a priori segmentation of the digital images used as stimuli in the experiment into ROIs: *iMap*. [It is worth noting that methods sharing similarities with *iMap* were introduced by various researchers before us (e.g., Barrington, Marks, Hsiao, & Cottrell, 2008; Bruce & Tsotsos, 2009; Buchan, Pare, & Munhall, 2007; Harding & Bloj, 2010; Henderson, 2003; Kita et al., 2010; Pomplun, Ritter, & Velichkovsky, 1996; Tatler, Wade, Kwan, Findlay, & Velichkovsky, 2010; Torralba, Oliva, Castelhana, & Henderson, 2006; Wooding, 2002). Here, we will briefly present the characteristics that few of these methods share with *iMap* and what differentiate them from it. Tatler, Baddeley & Gilchrist (2005) used a fixed grid with fixation counts in each cell and the Kullback–Leiber divergence (KL) in order to test differences in probability density functions. These authors did not weight the probability density functions with fixation durations. Moreover, because KL reports a single index for each comparison, Tatler et al. (2005) could not generate statistical fixation maps for single conditions (and their comparisons). Hence, in contrast with *iMap*, significant

differences between conditions could not be localised inside the stimulus space. Tatler (2007) used Gaussian smoothing in order to generate fixation maps. However, he did not weight the density functions by the fixation duration. Moreover, no statistical test was performed on the individual or difference fixation maps.

Note that using Gaussian smoothing and weighting by fixation durations is not new per se and heat-map representations have become very popular in the last few years (see for instance (e.g., Barrington, et al., 2008; Bruce & Tsotsos, 2009; Buchan, et al., 2007; Harding & Bloj, 2010; Henderson, 2003; Kita, et al., 2010; Tatler, et al., 2010; Torralba, et al., 2006; Wooding, 2002)). However, in most of the cases, heat-maps were just used for illustration purposes. More importantly, in comparison with *iMap*, no statistical test is performed on these fixation maps in order to locate the effects in the stimulus space. For instance Buchan et al. (2007) or Tatler et al. (2010) generated heat-maps for visualizing the eye-movement pattern but computed statistics using ROIs. Moreover, none of these approaches was implemented as a toolbox and is offering the numerous statistical and descriptive analyses we provide with *iMap*. For the whole stimulus space: number of fixations, total fixation duration, mean fixation duration, path length and mean saccade length. In the significant areas: Z-scored fixation durations (or number of fixations), effect sizes (Cohen's *d*). As well as, mean fixation duration, path length, total fixation duration and number of fixations for significant areas and the rest of the picture.

To the best of our knowledge, the method used by Leonards, Baddeley, Gilchrist, Troscianko, Ledda & Williamson (2007) is the most similar to *iMap*. These authors created fixation maps based on Gaussian kernels, generated difference maps and used robust statistics to compare conditions. The main advantage of *iMap* over their technique is its public availability, direct access to the parameters used and its ease of use. Another difference is that in Leonards et al. (2007), each fixation is replaced by an elongated Gaussian distribution

around this point (while *iMap* uses a circular Gaussian), and with a spread determined by the magnitude and angle of the saccade used to get to this location. This is a very interesting way to represent the distribution of landing positions depending of the direction and size of the saccade. However, in *iMap*, the Gaussian kernel is used to approximate a unique fixation location and not a distribution of fixations. In this sense, we think that using a circular (not elongated) Gaussian kernel allows us to keep as much as possible an assumption-free approach. In addition, one could argue that representing the direction of the fixation with elongated Gaussian is valid, but only for the few milliseconds following the arrival saccade arrival. Thereafter, a representation more close to the foveal projection (circular) might be more appropriate. But, similarly to *iMap*, this approach raises also novel questions: for instance, for how long an elongated representation is the most appropriate way to describe the data (20ms? 40 ms? etc.), is the shape of the Gaussian dependent of the task, background information, etc. at hand? Finally, it is also worth noting that although the direction of the saccade impacts on the landing distribution, the average across saccade directions reveals a nearly circular Gaussian distribution as used in *iMap* (see Figure 7 in Nuthmann & Henderson, 2010).

Other authors used slightly different data-driven approaches. We would like to particularly mention the Scanmatch toolbox from Cristino, Mathôt, Theeuwes & Gilchrist (2010), the scanpath similarity measure by Jarodzka, Holmqvist & Nystrom (2010) and the approach used by Mannan, Kennard & Husain (2009). Interestingly, Cristino et al. (2010) and Jarodzka et al. (2010) used similar methods: the Needleman–Wunsch algorithm for the first ones and the Levenshtein distance for the second ones (see also Harding & Bloj, 2010, for a use of this method). A key advantage of Cristino et al. (2010) is to provide an implemented toolbox, making such method available for researchers that are not expert in programming.

These previous above approaches aim at describing and comparing sequences of fixations, a depiction that *iMap* does not provide. However, both these techniques return a single number reflecting how similar some fixation sequences are. Thus, they do not allow, like *iMap*, to visualize and test statistically which spatial areas are fixated significantly longer and which areas show significant differences between two datasets. Note that Tatler et al. (2005) also investigated temporal sequences of eye movements by computing KL for each specific fixation. This strategy is also possible with *iMap*, individual and difference maps can be computed for single, successive fixations in order examine temporal characteristics of the oculomotor behavior.

Carmi & Itti (2006) presented and discussed various metrics for comparing fixation distributions. The approaches they present are particularly interesting for the question they explore: quantifying the agreement between human attentional selection and attention-priority maps. We will not detail these metrics here because none of them allows, as in *iMap*, to visualize the effects location in the stimulus space. The same limitation applies to the Voronoi diagrams that Over, Hooge and Erkelens (2006) used to provide a quantitative measure of the uniformity of fixation densities.

[]

To sum up, despite some similarities between those previous approaches, *iMap* remains an original, complementary tool for analysing eye movement data. The main difference between *iMap* and the methods previously cited above is that *iMap* provides an implemented toolbox that allows users to share an identical implementation of this technique and, above all, to compute robust statistical analyses. *iMap* generates fixation distributions smoothed with Gaussian kernels, transforming 2D fixation maps, uniquely based on the fixation coordinate location in x, y dimensional space, into a 3D fixation landscape, with z

reporting the intensity of the fixation (number of fixations or fixations weighted by their durations – see Figure 2).]

[Eye movements do not provide unequivocal evidence on the measure of the visual information being used by observers (Posner, 1980). Hence, it is important to stress that iMap generates statistical fixation maps and not the so-called attentional maps (e.g., <http://www.attentiontool.com>; <http://eyequant.com>). The Gaussian kernel is a variable that can be used to characterize both the visual information that can be sampled for a given fixation and/or the potential error due to the eye-tracker. This choice offers the flexibility to the user to set the kernel size according to specific hypothesis, material, population, task, equipment or presentation conditions. Researchers who would like to adopt an assumption-free approach can set the kernel to a minimum value corresponding to the accuracy of the eye-tracker. This strategy was the one used in our previous articles where the kernel size was set to 10 pixels corresponding to 0.5 degree of visual angle (accuracy of the Eyelink 1000, desktop) in our set-up. This flexibility allows researchers to adapt the kernel size to their equipment but to the participant as well. Hence, it is possible to set a specific kernel for each individual depending on the eye tracker accuracy, as measured during the calibration procedure.]

iMap generates fixation maps for each single participant and every visual stimulus, resulting on an average fixation map. The *individual* fixation maps can then be averaged together, resulting in a *group* fixation map (Figure 2).

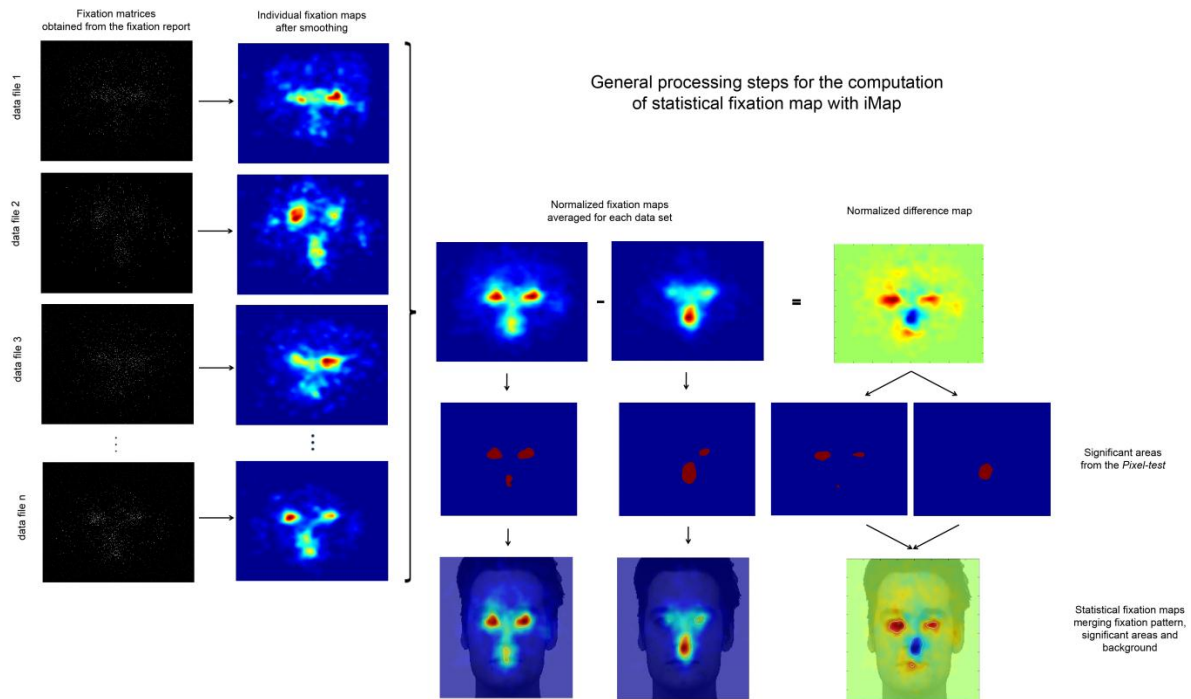


Figure 2. General processing steps for the computation of statistical fixation map with *iMap* (adapted from Caldara, et al. (2010)). Individual fixations maps are smoothed by convoluting a Gaussian kernel on each fixation. The resulting fixation maps for all the trials are then averaged to results in a *single* fixation map per condition. The differential fixation map highlights significant eye movement biases. The significant areas are determined using the *Pixel* test (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005). Finally, the statistical fixation maps are produced merging the fixation patterns, the areas significantly fixated above chance level and background.

But the critical value and the key innovative feature of the *iMap* technique relies on the ability to statistically compare fixation maps, with an approach taking into account the problem of the multiple comparisons generated by the pixel space. To the best of our knowledge, *iMap* is the first freely available technique integrating robust statistics in order to generate unbiased data-driven *statistical* fixation maps from eye movements. *iMap* corrects for multiple comparisons, quantifies the effect size of the statistical differences and also provides descriptive measures routinely used in eye movement research (i.e., number of fixations, average fixation duration, fixation scanpath length and total fixation duration).

The processing steps of the *iMap* method, its rationale and logic, were very much inspired by methods used in functional Magnetic Resonance Imaging (fMRI). The 3D fixation maps created with *iMap* can be considered as ‘activation maps’ of eye movement data. In fMRI studies, brain activations are represented by 3D maps containing many voxels (i.e., 3D volumetric pixel elements used to visualize and analyze fMRI data) over time (which is considered the fourth dimension). It is common practice to spatially smooth fMRI data by applying Gaussian kernels prior to the statistical comparison of brain activations across conditions (i.e., comparing the neural responses for processing faces to houses). Firstly, the smoothing procedure improves anatomical inter-subject variability. Secondly, it increases signal to noise ratio, by reducing random noise in individual voxels (Smith, 2003). Thirdly, this procedure ensures that the assumptions of Random Field Theory (RFT – see below), commonly used to correct for multiple-comparisons, are met (Worsley & Friston, 1995). Conventionally, fMRI results rely on massive univariate statistics testing for the effect of interest in each brain voxel, which therefore results in a large number of statistical comparisons, increasing the likelihood of *Type I* errors. Hence, in fMRI, the results need to be corrected for the multiple comparisons.

The statistical comparison of the 3D fixation maps generated with *iMap* share the very same problem, as the 3D fixation maps contain thousands of pixels and, therefore, generate the same large amount of statistical comparisons. In addition, similarly to voxels in the fMRI space, pixels are not statistically independent. Data for a particular pixel tend to be similar to the nearby pixels. RFT (Adler, 1981) is a recent branch of mathematics that has been implemented in statistics to overcome this major limitation. RFT has been adapted and used to define theoretical thresholds for smooth statistical maps in fMRI (Worsley et al., 1996). The RFT is based on two main processing stages. First, it relies on the estimation of the smoothness (spatial correlation) of the statistical maps. Then, it uses the smoothness values to

determine the expected *Euler characteristic* at different thresholds. This procedure estimates the threshold at which 5% of equivalent statistical maps are expected to arise under the null hypothesis. *iMap* relies on spatially normalized smoothed data which therefore satisfy the formal constraints of the RFT used in fMRI. More precisely, it applies the statistical *Pixel test* of the Stat4Ci toolbox (Chauvin, et al., 2005) that has been developed and validated for analyzing smooth classification images. The *Pixel test* sensitivity depends on the number of comparisons performed, which is represented here by the size of the search space (i.e., the size of the digital images). The default search space for *iMap* is the entire stimulus but a specific search space size can be specified. For instance, one could consider that the background of a picture with a face is not influencing eye movements during face processing and therefore reduce the search space to pixels belonging to the face only. To have a better understanding of some limitations in using *iMap*, this and other caveats of the approach will be addressed in the discussion, after providing formal knowledge of the technique and concrete examples.

[]

iMap has been already used and successfully validated in a series of eye movement studies (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Caldara, et al., 2010; Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Kelly et al., submitted; Kelly, Miellel, & Caldara, 2010; Miellel, et al., 2010; Rodger, Kelly, Blais, & Caldara, in press). To illustrate the functionality and flexibility of *iMap*, we provide examples from real eye movement data and diverse statistical comparisons (i.e., across different groups of observers, same observers but different tasks, etc.). Importantly, *iMap* has been coded with Matlab; the code is fully editable and is freely available to download and use. It is worth noting that we plan to continuously improve and update the *iMap* code and to also add *plug-ins* to the toolbox in the future (i.e., a *plug-in* to generate 3D fixation map movies or dynamic statistical tests over time). The relevant

information and files will always be freely available to download and use online. Finally, iMap can be used on data acquired with any eye-tracker and pre-processing software that provides a fixation report which includes the coordinates and duration of each fixation, as well as an item number.

Methods

Installation and credits

iMap and the supporting functions (CiVol.m, HalfMax.m, exportfig.m and stat_threshold.m) have to be copied in the same folder as the input data files.

Exportfig was written by Ben Hinkle, 2001 (bhinkle@mathworks.com) and can be downloaded at: <http://www.mathworks.com/matlabcentral/fileexchange/727>

CiVol and HalfMax are part of the Stat4Ci toolbox, which allows performing the Pixel and the Cluster tests, both based on Random Field Theory. The Stat4Ci toolbox is free and can be downloaded at:

<http://www.mapageweb.umontreal.ca/gosselif/basic%20Stat4Ci%20tools/>

If you use the statistical functions of the Stat4Ci called with iMap (i.e., *Pixel* or *Cluster test*), please cite Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M. & Gosselin, F. (2004). A sensitive statistical test for smooth classification images. *Journal of Vision*, 5, 659–667.

The stat_threshold function was written by Keith Worsley for the fmristat toolbox which is free to download at: <http://www.math.mcgill.ca/~keith/fmristat>

An alternative to copying some of the supporting functions into the data folder is to download the Stat4Ci and fmristat toolboxes and to add them to the Matlab path.

The Western Caucasian face images used in the examples of the toolbox belongs to the KDEF (Lundqvist, Flykt, & Öhman, 1998) face database.

Creating the input matrices

iMap can be used with eye movement data collected with any eye tracker. First, the data needs to be pre-processed in order to determine saccades and fixations. This filtering process should be feasible with the majority of the analysis software provided with eye-trackers, or alternatively, with a saccade detection algorithm based on the eye velocity.

The file resulting from this pre-processing will be used as input data for the *iMap* function. The input data is a set of matrices with a single fixation per line. The only data required are the coordinates and duration of the fixations, and the item numbers. The order of the columns has no importance, as they have to be specified in the *iMap* function. Any other column can be used for specifying experimental conditions. A specific input data matrix has to be created for each participant and/or condition. The input files used by the *iMap* function are Matlab .mat files called data1.mat, data2.mat,...). The matrix in each of the files is called "summary". The matrices and files can be made from any .txt file (e.g. fixation report from EyeLink® Data Viewer). Some of the examples show how to create such input data matrices.

Running the iMap function

iMap can then be used by calling a single function including a set of parameters. The general format of the function is:

imap (xSize, ySize, columnx, columny, columnduration, columnitem, dataset1, dataset2, standard deviation, maptype, firstfix, backgroundfile, specificfix, searchspace), with

- 1- xSize and ySize: stimulus size in pixels (e.g. 382, 390)
- 2- columnx, columny, columnduration, columnitem: specify the column number for x, y coordinates, fixation durations and item number. This allows flexible data format.
- 3- datasets 1 and 2: specify the data .mat files that will be tested/compared. For example [1:20], [21:40] to compare data1 to data20 with data 21 to data40. The second data set is

optional; this field has to be left empty if only one dataset is tested. If only one dataset is tested, iMap produces a statistical map and eye-tracking indexes for this dataset. If two datasets are specified, iMap provides the statistical maps and eye-tracking indexes for both dataset and the difference map and indexes.

4- standard deviation: Standard deviation in pixels of the Gaussian kernel used for smoothing the data. The default value is 10 pixels. Specifying a value is necessary in order to specify the subsequent variables of the function. The empty square brackets [] setting, will use the default value.

5- matype: 1 for fixation duration maps, 2 for number of fixations maps. The default value is 1.

6- firstfix: This option discards the first fixation of each trial. This is particularly useful if the stimuli are centred and a central fixation cross is presented before the trials. 1 (default option) keeps all the fixations, 2 ignores the first fixation of each trial.

7- backgroundfile: e.g. 'facebackground.tif'. This option allows adding a background picture to the statistical fixation maps. The value is optional and has to be set to 0 or [] in order to specify the subsequent variables.

8- specificfix: To select one or several specific fixations. e.g. [3 3] or [1 3]. This value is optional.

9- searchspace: By default the size of the stimulus, xSize * ySize. The search space size can be specified by indicating directly the number of pixels it contains or by using a black and white picture (e.g. 'facemask.tif') where the black mask indicates the search space.

Importantly, after launching the function a map will stay on the screen, maximize it then click on the top-left corner then bottom-right then "enter" (see figure 3):

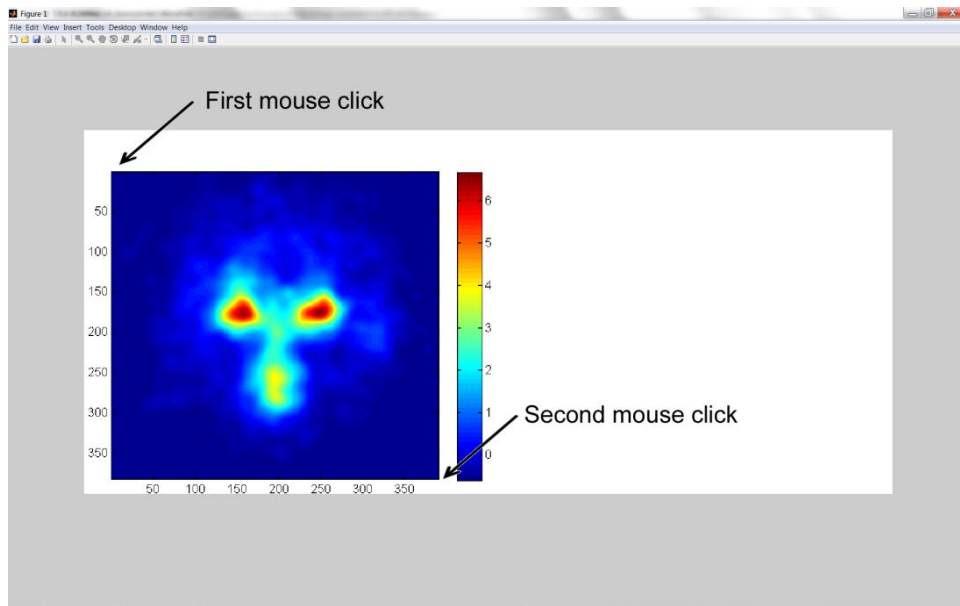


Figure 3. Validation of the fixation map area when using *iMap*.

Output

iMap creates .tif pictures of the single and difference fixation maps called `dataset1picedge.tif`, `dataset2picedge.tif` and `diffpicedge.tif` respectively. The maps can be merged with a background picture. It displays the significant areas based on a Pixel-test. It also creates .tif pictures with the scales of the Z-scored fixation measures, called `dataset1map.tif`, `dataset2map.tif` and `Zdiffmap.tif` respectively.

iMap generates .txt files with global eye-tracking measures for both datasets (called `eyebasicdataset1.txt` and `eyebasicdataset2.txt`). The columns are: the number of fixations, the total fixation duration (seconds), the mean fixation duration (seconds), the path length (pixels) and the mean saccade length (pixels). The lines correspond to the raw data files (participants, sessions). *iMap* also creates a text file called `Zscore.txt` that includes the mean Zscores in the significant area for (respective columns) the dataset 1, dataset 2, dataset 1 in the area 1 and area 2 (areas in which the fixation durations are significantly longer for dataset 1 and 2 respectively), dataset 2 in the area 1 and area 2.

iMap also produces a .txt file with the Cohen's d values (Cohen, 1988) between both datasets for area 1 and 2 (areas in which the fixation durations are significantly longer for dataset 1 and 2 respectively). The file is called cohend.txt. Finally, imap creates .txt files with the eye-tracking data in both the significant areas and the rest of the picture. The files are called eyeareadataset1.txt and eyeareadataset2.txt and are organised the following way: mean fixation duration for area 1 then for area 2 then for the rest of the picture. Path length, total fixation duration and number of fixations are also organised with the same logic.

Examples

For convenience, iMap and the supporting functions have been copied to each of the example folders.

Example 1:

Example 1 uses a subset of data from Caldara, Zhou & Miellat (2010). Putting Culture Under the 'Spotlight' Reveals Universal Information Use for Face Recognition. PLoS ONE 5(3): e9708. doi:10.1371/journal.pone.0009708

In this experiment, East-Asian (EA) and Western-Caucasian (WC) participants performed an old-new task on EA and WC faces. The stimuli came from the KDEF (Lundqvist, et al., 1998) and the Asian Face Image Database (Bang, Kim, & Choi, 2001). The presentation was gaze-contingent with a 2°, 5° or 8° Gaussian aperture around the fixation location. For this example, the aperture size is 8°. There was a central fixation cross before each trial then the 382x390 stimulus was randomly placed on an 800x600 screen. The eye position was recorded every 8ms with Matlab. We then extracted fixations and saccades (with

a custom-made Matlab script), by using the same filter parameters as the EyeLink software: saccade velocity threshold = 30°/sec; saccade acceleration threshold = 9500°/sec.

The data in this example are in .mat files (called data1.mat, data2.mat,...), the matrices are called "summary".

The 1st contrast aims at comparing the respective eye movement strategies of Western Caucasian [2 3 8 12 13 18] versus East Asian [22 23 24 32 33 34] observers deployed when learning human faces (WC and EA faces stimuli for both groups). The values for matype and firstfix were 1, so the fixation duration maps were generated and the first fixation of each trial was included in the analysis (note that the position of the stimulus was randomized on the screen). The *iMap* function can then be executed by typing:

```
imap (382, 390, 6, 7, 5, 1, [2 3 8 12 13 18], [22 23 24 32 33 34], 10, 1, 1,
'facebackground.tif')
```

The statistical fixation maps produced are shown in figure 4:

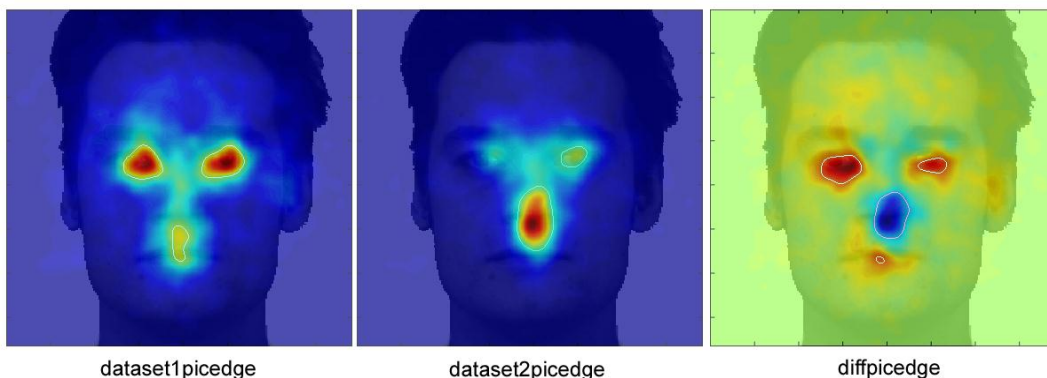


Figure 4. Statistical fixation maps for the first and second data sets, and for their difference.

This example shows the presence of significant fixation biases across the two group of observers (i.e., areas delimited with a white border). WC observers showed a fixation bias towards the eyes and mouth (dataset1picedge, red color in the difference map), whereas EA

observers showed a fixation bias towards the center of the face (dataset2picedge, blue color in the difference map).

The scaling can be obtained with the dataset1map.tif, dataset2map.tif and Zdiffmap.tif files.

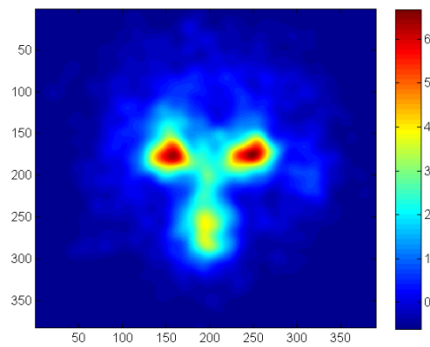


Figure 5. dataset1map.tif with the scale.

The numerical outputs produced by the analysis are reported below. Here, for simplicity, we report only the global eye-tracking measures and the measures according to significant areas for the dataset1.

number of fixations	total fixation duration	mean fixation duration	path length	mean saccade length
29.09677	7.483355	0.2343271	1539.18	48.12055
27.87097	5.932745	0.2136969	1699.69	61.15729
27.67742	7.76056	0.2542954	1674.231	54.56829
28.41935	7.174361	0.2308422	1353.442	43.39324
27.87097	4.96008	0.1603119	2130.377	69.2846
27.32258	7.850284	0.2606777	1726.404	57.04416

Table 1. Eyebasicdataset1.txt, averages of global eye-tracking measures for the dataset 1

mean fixation duration						total fixation duration					
(sec.)			path length (pixels)			(sec.)			number of fixations		
area 1	area 2	rest	area 1	area 2	rest	area 1	area 2	rest	area 1	area 2	rest
0.3178	0.2506	0.2521	135	47	65	0.8901	0.2753	7.1256	2.79	1.00	28.46
0.1881	0.2297	0.2128	63	117	90	0.1566	0.4580	5.3216	0.87	1.87	25.16
0.3166	0.2829	0.2778	300	71	89	1.4285	0.3257	6.8447	4.75	1.18	24.75
0.2904	0.2448	0.2540	108	53	82	0.7069	0.3226	6.9384	2.50	1.39	27.61
0.1808	0.1693	0.1775	89	59	63	0.2640	0.1710	5.0603	1.46	1.00	28.43
0.3323	0.3189	0.2759	266	113	86	1.5321	0.6282	6.5381	4.64	1.96	23.68

Table 2. Eyereadataset1.txt, averages of eye-tracking measures for significant areas, dataset 1

single maps			difference map			
		dataset1 -	dataset1 -	dataset1 -	dataset1 -	
dataset1	dataset2	area1	area2	area1	area2	
4.483855	4.909454	4.646505	2.261218	1.699236	5.669345	

Table 3. Z-score.txt. Z-scored fixation durations in significant areas

cohen's d	cohen's d
area1	area2
1.575624	-1.697319

Table 4. cohend.txt. Effect sizes in the significant areas of the difference map

For this particular task, it was appropriate to specify only the face area as search space (excluding the white background). Indeed, if the *default* search space contains a large amount

of pixels that are never fixated (because there is no visual information for instance), the Z_{crit} can be too low to accurately capture the sensitivity of eye movement patterns. Adjusting the search space to an appropriate level of analysis can be done by using a mask to limiting the search to pixels containing information (a face mask here):

```
imap ( 382, 390, 6, 7, 5, 1, [2 3 8 12 13 18], [22 23 24 32 33 34], 10, 1, 1,
'facebackground.tif', [], 'facemask.tif')
```

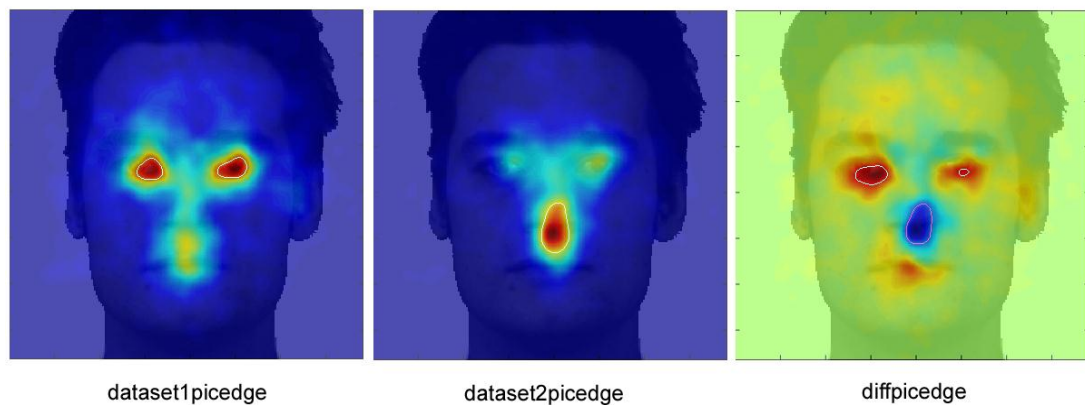


Figure 6. Statistical fixation maps for the first and second data sets, and for the difference when using a mask limiting the search space

Reducing the search space increases the threshold hence decreasing the sensitivity, as highlighted by the *iMap* analysis reported in Figure 6.

Example 2:

This example uses a subset of data from Mielle S., Zhou X., He L., Rodger H. & Caldara R. (2010). Investigating cultural diversity for extrafoveal information use in visual scenes. *Journal of Vision* Vol.10(6) pp 21.

<http://www.journalofvision.org/content/10/6/21.full>

In this experiment, the participants had to detect and identify an animal in a natural visual scene (full-screen, colour pictures). The two main manipulations were the size of the

target and the size of a gaze-contingent artificial central scotoma (*Blindspot*). The target and the Blindspot sizes could be 0° (natural vision), 2°, 5° or 8° of visual angle. The raw eye-tracking data were recorded in Matlab. The pre-processing was done with the velocity based saccade detection algorithm described above.

The target position was randomly distributed in the scene. In order to make the fixation maps, either trials can be considered individually (see an example with *singlescenes.m*) or fixation positions can be normalized relatively to the target position, by creating a new fixation space where all the targets are centered in the middle of the screen (see an example with *normalizedscenes.m*).

singlescenes.m shows how to generate data1... datan files containing the summary matrix with gaze coordinates, fixation durations and item number. Here, we created such files only for the no-*Blindspot* (0°) and 5°-target conditions. Moreover, because there is no spatial normalization in this example, we selected only the fixations corresponding to a specific item (here, item 49). In this example, only one dataset (including 10 participants) is considered and the first fixation of each trial was excluded (there was a central fixation cross before each trial and the stimuli were covering the full screen). Hence, the *iMap* function can be executed by typing the following sequence:

```
imap(600, 800, 1, 2, 3, 4, [1:10], [], 10, 1, 2, '5deg_9.tif')
```

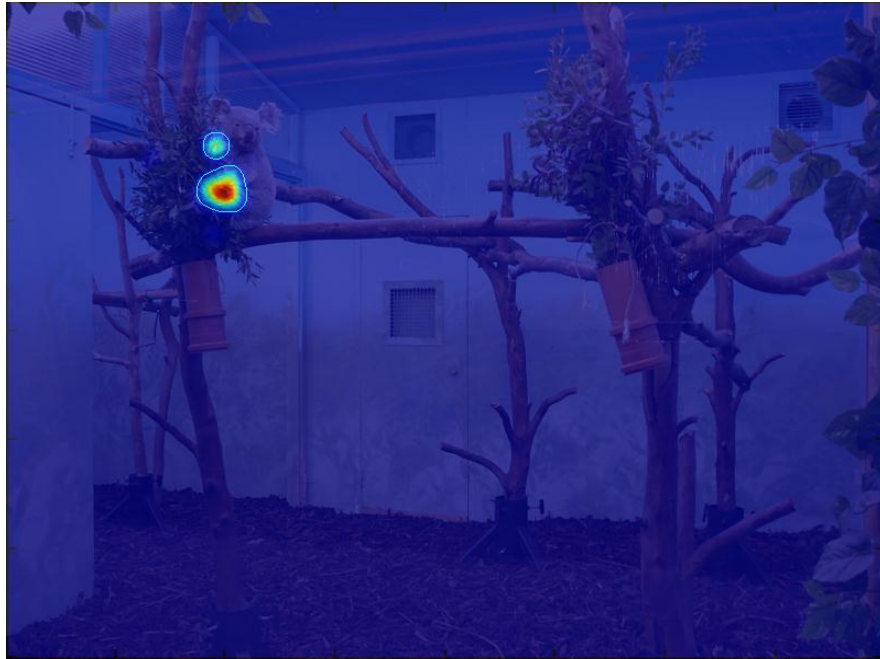



Figure 7. Statistical fixation map (fixation durations) for a specific item, the first fixation of each trial was excluded and the search space cover the entire stimulus/screen.

normalizedscenes.m is constructed in a similar way as *singlescenes.m*, beside the fact that there is no need to filter the data for a specific trial as the target positions are spatially normalized. Here, there is no background as different stimuli are considered. The *iMap* function can be executed by typing the following sequence:

```
imap(600, 800, 1, 2, 3, 4, [1:10])
```

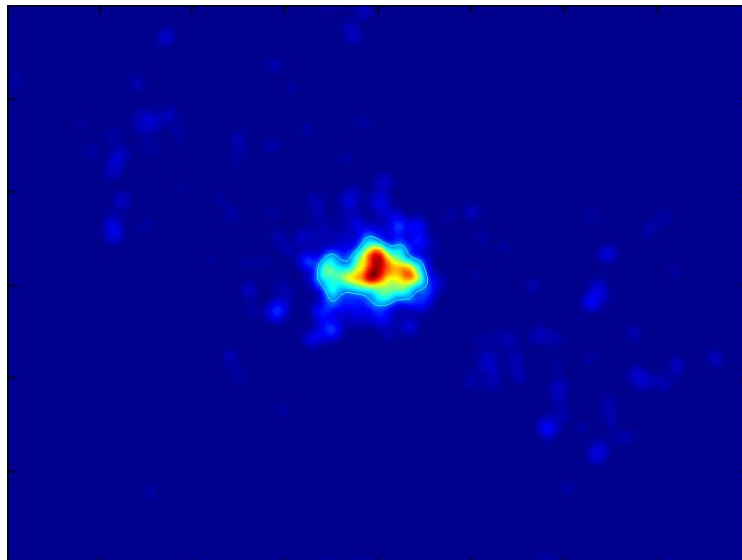


Figure 8. Statistical fixation map for spatially normalized items

Below is an example of a selection of specific fixations (fixations 6 to 15 of each trial, late fixations):

```
imap(600, 800, 1, 2, 3, 4, [1:10], [], 10, 1, 2, [], [6:15])
```

The syntax `imap(600, 800, 1, 2, 3, 4, [1:10], [], 10, 1, 2, 0, [6:15])` is also accepted.

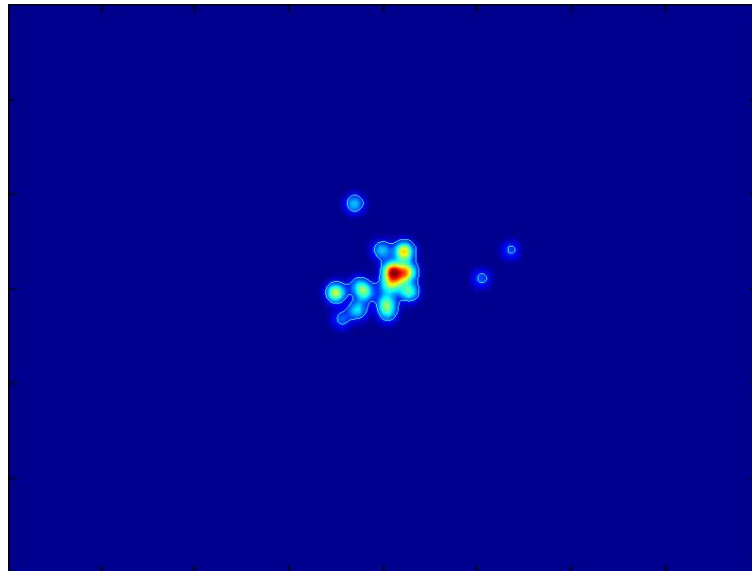


Figure 9. Statistical fixation map for spatially normalized items and late fixations

This example reveals less surrounding fixations due to the image exploration. In contrast, only a central hotspot is present showing that most of the ‘late’ fixations are on the target. This example also shows that it is possible to extract individual or a series of fixations for the eye movement analysis with *iMap*.

Example 3:

This example uses data collected during an experiment using eye movements and a memory task similar to the one described in Harkin, B., & Kessler, K., (2009). How Checking Breeds Doubt: Reduced Performance in a Simple Working Memory Task. *Behaviour Research and Therapy*, 47, 504-512.

The experiment involved learning letters position on a 2x3 grid, testing high vs. low-checking participants and the presentation of correct versus incorrect probes (see figure 10).

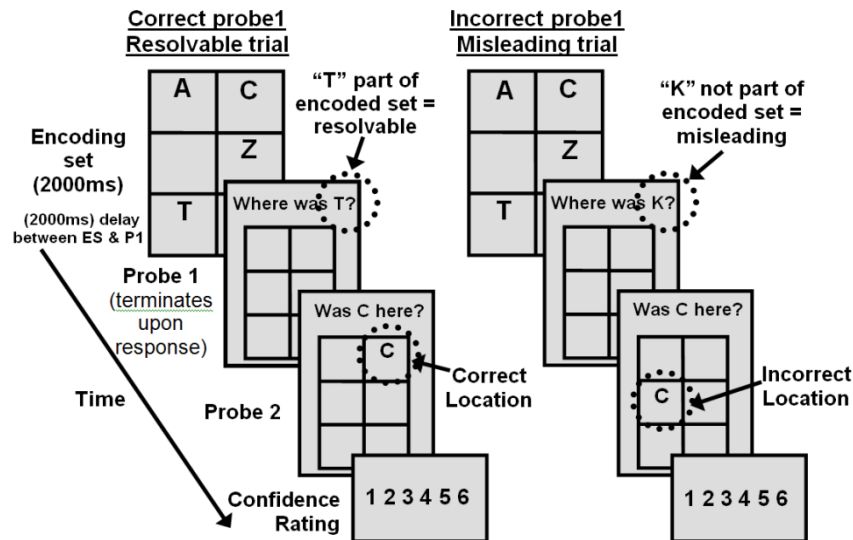


Figure 10. Presentation of the stimuli used and the time course of the trials.

The experiment was presented with E-Prime. The raw data were recorded in SR-Research edf format. The data were then pre-processed using SR-Research DataViewer and the fixation report was exported in text format. *memorytask.m* prepares the data for the analysis and run *iMap*. The screen-based coordinates are also centered on the stimulus.

The preparation code (*memorytask.m*) allows considering specific conditions. We can specify for correct or incorrect probe 1 (*corP1* and *incorP1*) and several time periods (period $1 < 2$ sec., $2 \text{ sec.} < \text{period } 2 < 4$ sec., $4 \text{ sec.} < \text{period } 3$). The fixation maps reveal specific patterns for each time period.

Figure 11 shows the fixation pattern for the learning stage (period), regardless of the probe correctness or the participant group (note that a specific analysis showed no effect of these factors during learning). We used the following function:

```
imap(396, 288, 1, 2, 3, 4, [lowcheck highcheck], [], 10, 1, 1)
```

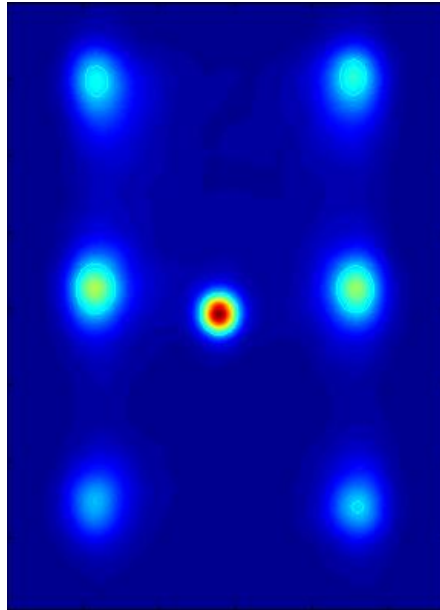


Figure 11. Fixation map of the learning stage calculated across all participants and probe conditions.

These data clearly show that the central fixation cross appearing before the beginning of the trial impacts on the fixation pattern. In such experimental situations (with no randomization of stimulus location and identical first fixation location for all the trials), it is recommended that the the first fixation is excluded. Figure 12 presents the same analysis when excluding the first fixation, by using the following parameters in the input of the *iMap* function:

```
imap(396, 288, 1, 2, 3, 4, [lowcheck highcheck], [], 10, 1, 2)
```

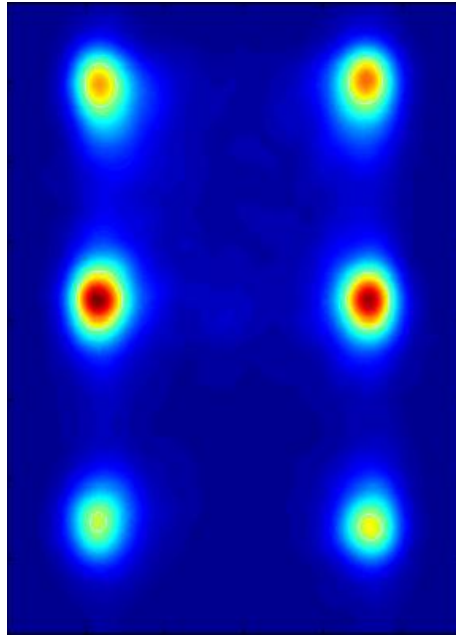


Figure 12. Fixation map of the learning stage calculated across all participants and probe conditions without the first fixation.

The statistical fixation map reveals significant hotspots on each of the 6 positions where the letters could appear. It also shows an upper-field bias.

Figure 13 shows the fixation pattern during the delay (period 2), regardless of the probe correctness or the participant group.

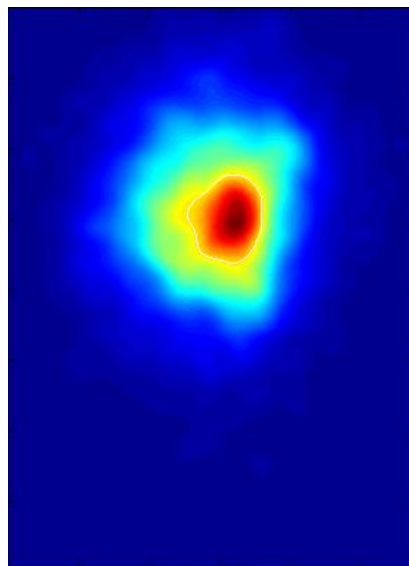


Figure 13. Statistical fixation map for the delay period (empty screen)

During the delay, no information was presented on the screen. The statistical fixation maps revealed central fixations with an upper-field bias and no effect of the probe or the group of participant.

Interestingly, the fixation maps were different for low- versus high-checkers in the third time period when the probe was incorrect. During the third period, the empty grid was presented and the participant had to indicate the location of the probe. The hotspots, at the top of the fixation maps, indicate that the participants were gazing at the instruction (probe, the letter they had to localize). Looking at the incorrect probe is sufficient for a response from the low-checkers. In contrast, the high-checkers verify on the grid (which is empty during this period) before answering.

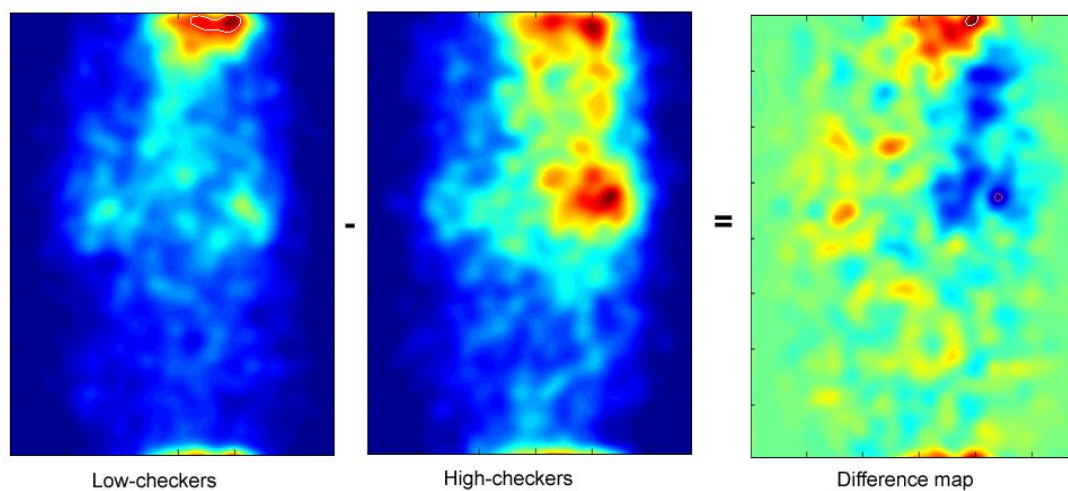


Figure 14. Statistical fixation maps for low- versus high-checkers during the third period in the incorrect-probe condition.

Discussion

We have developed an alternative method to analyze eye movement data: *iMap*. As previous approaches (e.g., Barrington, et al., 2008; Bruce & Tsotsos, 2009; Buchan, et al., 2007; Harding & Bloj, 2010; Henderson, 2003; Kita, et al., 2010; Pomplun, et al., 1996; Tatler, et al., 2010; Torralba, et al., 2006; Wooding, 2002), this method does not rely on the *subjective* definition of ROIs; it simply does not require the use of ROIs. *iMap* offers some advantages compared to previous methods. Firstly, it relies on robust statistics to assess the significance of the effects. Secondly, it is coded as an editable toolbox for Matlab, freely available for use.

To illustrate the functionality and flexibility of the toolbox, we have provided three examples. The results from those examples, coupled with the results from our previous work on face (Blais, et al., 2008; Caldara, et al., 2010; Jack, et al., 2009; Kelly, et al., submitted; Kelly, et al., 2010; Rodger, et al., in press) and scene processing (Miellet, et al., 2010) that are largely consistent with the Western Caucasian eye movement literature and the literature on East Asian observers (Kita, et al., 2010), show that the toolbox effectively captures eye movement sensitivity for the tasks at hand. In the first dataset, we initially compared the fixation strategies deployed by Western Caucasian (WC) and East Asian (EA) observers while learning WC and EA faces. This comparison resulted in significant fixation biases across observers. WC observers fixated the eye region more than EA observers, whereas EA observers fixated the central part of the face more in comparison with WC observers. Additional analyses also revealed similar fixation patterns for both type of stimuli (WC versus EA faces) or correct versus incorrect face recognition (see Caldara, et al., 2010). In the second dataset, observers had to detect and identify an animal in a natural visual scene. We showed *iMap* analyses for the natural digital images and digital images normalized (centred) for the position of the animal. This analysis showed significant fixation hotspots on the

search target. Additional analysis also revealed fixations on the targets despite large *Blindspots* and similar fixation patterns for EA versus WC participants (see Mielle, et al., 2010), along with the flexibility of using a subset of fixations for generating the statistical maps. Finally, in the last dataset, we compared High- and Low-checkers observers in a memory task. This analysis showed that High- and Low-checkers deploy different strategies when confronted with ambiguous/erroneous information.

iMap was inspired by *fMRI* methods and it suffers for very similar caveats, which we are going to address in turn: the choice of the width of the standard deviation of the Gaussian kernel used during the smoothing procedure and the normalization.

The first parameter is clearly dependent on the experimental stimuli used in the eye movement study and the task at hand. In neuroimaging, this choice is perhaps easier, as the width of the Gaussian should not be larger than the brain region of interest, for instance: 3 to 6 mm for the Full Width Half Maximum for a small region such as the Fusiform Face Area (e.g., Caldara & Seghier, 2009; Caldara et al., 2006; Rossion et al., 2003; Schiltz et al., 2006), or up to 12mm for the Full Width Half Maximum for a larger region such as the insula (see Mutschler et al., 2007). The logic is similar for eye movement analyses. In our previous work, we were very careful to not “over-smooth” our data. We have used Gaussian kernels with a standard deviation covering approximately 0.5° of visual angle (Blais, et al., 2008; Caldara, et al., 2010; Jack, et al., 2009; Kelly, et al., submitted; Kelly, et al., 2010; Mielle, et al., 2010; Rodger, et al., in press), which is roughly the size covering a fourth of the fovea (Hood & Finkelstein, 1986). We thought this was a sensible choice for the question we aimed to address: investigating cultural diversity in face processing. However, this could not be considered as the optimal parameter for analyzing *any* eye movement task. For instance, let's assume we would like to investigate the role of the pupil size in the evaluation of attractiveness for a series of human face stimuli. In this case, it would be necessary to

significantly decrease the standard deviation of the Gaussian kernel in order to capture subtle differences in the fixation patterns falling in the pupil/eye region. As for neuroimaging, there is no a governing rule to define the size of the standard deviation of the Gaussian kernel. The experimenters need to evaluate and back up their choice with existing knowledge in the literature, and also justify their choices in respect to the stimuli, the equipment and the task used in the experiment.

Similar to neuroimaging data that are recorded from brains of different participants, *iMap* requires a normalized space to perform statistical analyses. Without entering into the details, in *fMRI* there are several approaches used to normalize the human brains of different participants. The most commonly used approaches involve the realignment of the brains into the Talairach (Talairach & Tournoux, 1988) or the Montreal Neurological Institute (MNI) standard template spaces. This process ensures that the comparison of voxels across participants is valid. Similarly, *iMap* requires that the fixation landscapes created for a particular condition are built on a homogenous space, such as a given fixation for one item is fully comparable to a fixation on the same location in another item. Therefore, to meet this requirement, the faces we used in our previous studies (Blais, et al., 2008; Caldara, et al., 2010; Jack, et al., 2009; Kelly, et al., submitted; Kelly, et al., 2010; Rodger, et al., in press) and in the current examples, were normalized for their eye and mouth positions. In the example we provide for visual scenes, we demonstrated that is also possible to normalize natural scenes, by arbitrarily centring the object of interest in the middle of the scene (see Mielliet, et al., 2010); note that the task used here was to find and identify the animal. However, *iMap* does not prevent the analyses of a unique input space, as long as many eye movement samples are collected to ensure the statistical validity of the analysis.

It is worth noting that *iMap* has been developed to analyse *where* and *when* eye movements are performed by the observers. As illustrated by the previous examples, *iMap*

can extract (and compare) the fixation maps of each particular fixation (first, second, etc.). This descriptive analysis provides information on the time course of fixations. Recently, there have been various fruitful attempts to integrate these measures occurring over time and to extract the occurrence of *statistically* significant sequences in the scanpath used by the observers (e.g., Cristino, et al., 2010; Jack, et al., 2009).

To sum up, *iMap* can analyse eye movement data with a robust data-driven approach that generates statistical fixation maps. As with every novel method, we anticipate improvements in the near future arising from the feedback of potential users. We aim to keep the *iMap* method updated, and will freely provide new versions of the Matlab toolbox code online. We believe that various approaches and methods are necessary in any scientific discipline, allowing researchers to use the more appropriate method to answer the question at hand. We hope that users will help us on improving *iMap* and eventually on building bridges with other data-driven Matlab-based toolboxes for eye-movement analysis.

Acknowledgments

The authors were supported by The Economic and Social Research Council and Medical Research Council (ESRC/RES-060-25-0010). We would like to thank Helen Rodger for her comments on a previous version of the paper.

References

- Adler, R. J. (1981). *The Geometry of Random Fields*. New York: Wiley.
- Bang, S., Kim, D., & Choi, S. (2001). Asian Face Image Database. Postech, Korea.
- Barrington, L., Marks, T. K., Hsiao, J. H., & Cottrell, G. W. (2008). NIMBLE: a kernel density model of saccade-based visual memory. *J Vis*, 8(14), 17 11-14.
- Barton, J. J., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: the effects of familiarity, inversion, and morphing on scanning fixations. *Perception*, 35(8), 1089-1105.
- Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE*, 3(8), e3022.
- Bruce, N. D., & Tsotsos, J. K. (2009). Saliency, attention, and visual search: an information theoretic approach. *J Vis*, 9(3), 5 1-24.
- Buchan, J. N., Pare, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Soc Neurosci*, 2(1), 1-13.
- Buswell, G. T. (1935). *How People Look at Pictures: A Study of The Psychology of Perception in Art*. Chicago: The University of Chicago Press.
- Caldara, R., & Seghier, M. L. (2009). The Fusiform Face Area responds automatically to statistical regularities optimal for face categorization. *Hum Brain Mapp*, 30(5), 1615-1625.
- Caldara, R., Seghier, M. L., Rossion, B., Lazeyras, F., Michel, C., & Hauert, C. A. (2006). The fusiform face area is tuned for curvilinear patterns with more high-contrasted elements in the upper part. *Neuroimage*, 31(1), 313-319.

- Caldara, R., Zhou, X., & Mielle, S. (2010). Putting culture under the 'spotlight' reveals universal information use for face recognition. *PLoS ONE*, 5(3), e9708.
- Carmi, R., & Itti, L. (2006). The role of memory in guiding attention during natural vision. *J Vis*, 6(9), 898-914.
- Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *J Vis*, 5(9), 659-667.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cristino, F., Mathot, S., Theeuwes, J., & Gilchrist, I. D. (2010). ScanMatch: a novel method for comparing fixation sequences. *Behav Res Methods*, 42(3), 692-700.
- Dodge, R. (1916). Visual motor functions. *Psychological Bulletin*, 13, 421-427.
- Harding, G., & Bloj, M. (2010). Real and predicted influence of image manipulations on eye movements during scene recognition. *J Vis*, 10(2), 8 1-17.
- Harkin, B., & Kessler, K. (2009). How Checking Breeds Doubt: Reduced Performance in a Simple Working Memory Task. *Behaviour Research and Therapy*, 47, 504-512.
- Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends Cogn Sci*, 7(11), 498-504.
- Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Mem Cognit*, 33(1), 98-106.
- Hering, E. (1879). Über Muskelgerausche des Auges. *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften in Wien. Mathematisch-naturwissenschaftliche Klasse, Abt. III*(79), 137-154.
- Hood, D. C., & Finkelstein, M. A. (1986). *Sensitivity to light, Sensory processes and perception (Vol. 1)*. New York: Wiley.

- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Curr Biol*, *19*(18), 1543-1548.
- Jarodzka, H., Holmqvist, K., & Nyström, M. (2010). *A Vector-based, Multidimensional Scanpath Similarity Measure*. Paper presented at the Eye-tracking Research and Applications (ETRA).
- Javal, É. (1879). Essai sur la physiologie de la lecture. *Annales d'oculistique*, *82*, 242-253.
- Kelly, D. J., Jack, R. E., Mielle, S., De Luca, E., Foreman, K., & Caldara, R. (submitted). Social experience does not abolish cultural diversity in eye movements.
- Kelly, D. J., Mielle, S., & Caldara, R. (2010). Culture shapes eye movements for visually homogeneous objects. [Original Research Article]. *Frontiers in Perception Science*, *1*.
- Kita, Y., Gunji, A., Sakihara, K., Inagaki, M., Kaga, M., Nakagawa, E., et al. (2010). Scanning strategies do not modulate face identification: eye-tracking and near-infrared spectroscopy study. *PLoS ONE*, *5*(6), e11050.
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci*, *12*(5), 535-540.
- Lamare, M. (1892). Des mouvements des yeux dans la lecture. *Bulletin et Memoire de la Societe Francaise d'Ophthalmologie*(10), 354-364.
- Leonards, U., Baddeley, R., Gilchrist, I. D., Troscianko, T., Ledda, P., & Williamson, B. (2007). Mediaeval artists: masters in directing the observers' gaze. *Curr Biol*, *17*(1), R8-9.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). The karolinska directed emotional faces. Stockholm, Sweden: Karolinska Institute.
- Mannan, S. K., Kennard, C., & Husain, M. (2009). The role of visual salience in directing eye movements in visual object agnosia. *Curr Biol*, *19*(6), R247-248.

- Miellet, S., Zhou, X., He, L., Rodger, H., & Caldara, R. (2010). Investigating cultural diversity for extrafoveal information use in visual scenes. *Journal of Vision, 10*(6), 1-18.
- Mutschler, I., Schulze-Bonhage, A., Glauche, V., Demandt, E., Speck, O., & Ball, T. (2007). A rapid sound-action association effect in human insular cortex. *PLoS ONE, 2*(2), e259.
- Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *J Vis, 10*(8), 20.
- Orban de Xivry, J. J., Ramon, M., Lefevre, P., & Rossion, B. (2008). Reduced fixation on the upper area of personally familiar faces following acquired prosopagnosia. *J Neuropsychol, 2*(Pt 1), 245-268.
- Over, E. A., Hooge, I. T., & Erkelens, C. J. (2006). A quantitative measure for the uniformity of fixation density: The Voronoi method. *Behav Res Methods, 38*(2), 251-261.
- Pomplun, M., Ritter, H., & Velichkovsky, B. (1996). Disambiguating complex visual information: towards communication of personal views of a scene. *Perception, 25*(8), 931-948.
- Posner, M. I. (1980). Orienting of attention. *Q J Exp Psychol, 32*(1), 3-25.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychol Bull, 124*(3), 372-422.
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Q J Exp Psychol (Colchester), 62*(8), 1457-1506.
- Rodger, H., Kelly, D., Blais, C., & Caldara, R. (in press). Inverting faces does not abolish cultural diversity in eye movements.

- Rossion, B., Caldara, R., Seghier, M., Schuller, A. M., Lazeyras, F., & Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, *126*(Pt 11), 2381-2395.
- Schiltz, C., Sorger, B., Caldara, R., Ahmed, F., Mayer, E., Goebel, R., et al. (2006). Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus. *Cereb Cortex*, *16*(4), 574-586.
- Smith, S. M. (2003). *Preparing fMRI data for statistical analysis*. Oxford, UK: Oxford University Press.
- Talairach, J., & Tournoux, P. (1988). *Co-planar Stereotaxic Atlas of the Human Brain*. Stuttgart: Thieme.
- Tatler, B. W. (2007). The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. *J Vis*, *7*(14), 41-17.
- Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of fixation selection: effects of scale and time. *Vision Res*, *45*(5), 643-659.
- Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2010). Yarbus, eye movements, and vision. *I*(1), 7-27.
- Torralba, A., Oliva, A., Castelano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. *Psychol Rev*, *113*(4), 766-786.
- Van Gompel, R. P. G., Fischer, M. H., Murray, W. S., & Hill, R. L. (2007). *Eye movements: A window on mind and brain*. Oxford: Elsevier.
- Wade, N. J., Tatler, B. W., & Heller, D. (2003). Dodge-ing the issue: Dodge, Javal, Hering, and the measurement of saccades in eye-movement research. *Perception*, *32*(7), 793-804.

- Wooding, D. S. (2002). Eye movements of large populations: II. Deriving regions of interest, coverage, and similarity using fixation maps. *Behav Res Methods Instrum Comput*, 34(4), 518-528.
- Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited--again. *Neuroimage*, 2(3), 173-181.
- Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Hum Brain Mapp*, 4(1), 58-73.
- Yarbus, A. L. (1965). *Role of eye movements in the visual process*. Moscow: Nauka.
- Yuval-Greenberg, S., Tomer, O., Keren, A. S., Nelken, I., & Deouell, L. Y. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron*, 58(3), 429-441.