# *i*Map: a novel method for statistical fixation mapping of eye movement data

Roberto Caldara · Sébastien Miellet

**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 article, we introduce an alternative method for computing statistical fixation maps of eye movements—*i*Map—based on an approach inspired by methods used in functional magnetic resonance imaging. Importantly, *i*Map does not require the a priori segmentation of the experimental images into ROIs. With *i*Map, fixation data are first smoothed by convolving Gaussian kernels to generate three-dimensional fixation maps. This procedure embodies eyetracker accuracy, but the Gaussian kernel can also be flexibly set to represent acuity or attentional constraints. In addition, the smoothed fixation data generated by *i*Map conform to the assumptions of the robust statistical random field theory (RFT) approach, which is applied thereafter to assess significant fixation spots and differences across the three-dimensional 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 *i*Map, we provide sample analyses of real eye movement data from face, visual scene, and memory processing. The *i*Map MATLAB toolbox is editable and freely available for download online (www.unifr.ch/psycho/ibmlab/).

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, a description of the use of fixations to gather the information relevant to solving 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 for the scientific study of eye movements (see Wade, Tatler, & Heller, 2003).

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 similar oculomotor behavior, 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 do not randomly sample the visual input space, but are effective at solving problems in visual cognition. This work was then followed by a series of studies, leading in

R. Caldara (✉)
Department of Psychology, University of Fribourg,
Faucigny 2,
1700 Fribourg, Switzerland
e-mail: roberto.caldara@unifr.ch

S. Miellet
Centre for Cognitive Neuroimaging,
Institute of Neuroscience and Psychology, University of Glasgow,
Glasgow, Scotland

 Springer

particular to the seminal work of Yarbus (1965), which extensively showed similar findings for diverse visual objects. More importantly, Yarbus showed how top-down factors modulate the eye movement strategies deployed by observers to gather information from the very same picture, a discussion that attracted attention in the wider scientific community.[1] Since then, we have witnessed an explosion of eye movement studies on the processing of language and visual scenes (see Rayner, 1998, 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, ease, and affordability of eyetracking 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 have led 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 types of events characterizing *where*, *when*, and *how* the eyes gather information from the visual world—*saccades* (i.e., their latency, amplitude, direction, and occurrence over time) and *fixations* (i.e., their location and duration)—as well as many measures derived from these events, such as cumulative saccade length, pupil dilation, and so on. In order 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-interest (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, as well as high-level semantic a priori expectations 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 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 optimally represent the visual categories present in the visual input space (e.g., for face processing, the eye, 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 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, besides an *objective* definition of an ROI based on the human sclera, the remaining options used to define ROIs for the 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 regions of 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 of 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 (e.g., head, neck, hands), or would a unique shape defining the body be more appropriate? Should an 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, but under certain circumstances using ROIs might not be appropriate for thoroughly and effectively capturing eye movement behavior (see Fig. 1).

The more critical consequences of this drawback, however, 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 the qualitative and quantitative subjective evaluations of the experimenters, they engender natural variations across authors, which in some cases lead to difficulty in generalizing observations across studies. Note that other potential problems of using ROIs have also been

---

[1] Note that Yarbus (1965) recorded only a single observer for demonstrating task effects on eye movement patterns.
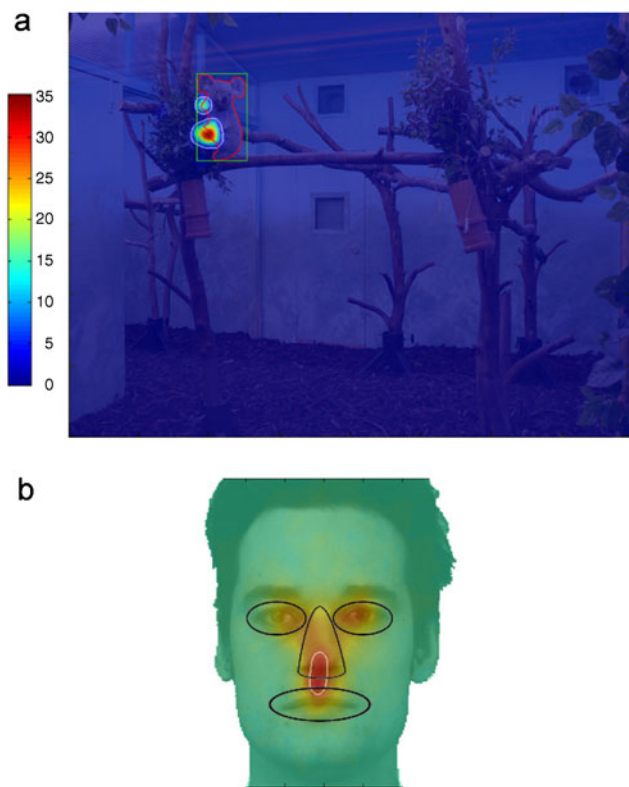
**Fig. 1** (**a**) Example extracted from a trial of the animal visual search task used in Miellet et al.'s (2010) study. The red contour represents an ROI based on the edges of the koala, a rule routinely used in the eye movement literature relying on ROIs. The white contours show areas of the visual scene that were fixated significantly above chance level using *i*Map. The center of gravity of the location driving the majority of fixations in this image is located outside the ROI. This perceptual bias of the observers would therefore be inaccurately reported by probability-of-fixation analyses based on ROIs. (**b**) Diagram illustrating the difficulty of defining a priori ROIs in face processing studies (data from Caldara et al., 2010). The white contour shows a significant area according to *i*Map, and blue contours show examples of ROIs as commonly used in the literature. In this example, it is difficult to objectively attribute the intermediary fixations to the nose or mouth regions

discussed in neuroimaging, particularly the dangers of circular analyses and "double dipping"—the use of the same dataset for selection and selective analysis (see Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). Therefore, the subjective variation of ROIs alone could potentially explain the absence of consistent effects across studies reported in the eye movement literature and points toward a methodological problem.

To overcome these limitations, we have introduced a novel, robust, data-driven technique that does not require the a priori segmentation of digital images used as stimuli into ROIs: *i*Map. *i*Map generates fixation maps for each single participant and every visual stimulus, resulting in an average fixation map. The *individual* fixation maps can then be averaged together, resulting in a *group* fixation map (Fig. 2).

The critical value and key innovative feature of the *i*Map technique, however, is 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, *i*Map is the first freely available technique integrating robust statistics in order to generate unbiased, data-driven *statistical* fixation maps from eye movements. *i*Map corrects for multiple comparisons, quantifies the effect size of the statistical differences, and provides descriptive measures routinely used in eye movement research (i.e., number of fixations, average fixation duration, fixation scan path length, and total fixation duration).

The processing steps of the *i*Map method, its rationale and logic, were very much inspired by methods used in functional magnetic resonance imaging (fMRI). The three-dimensional (3-D) fixation maps created with *i*Map can be considered as "activation maps" of eye movement data. In fMRI studies, brain activations are represented by 3-D maps containing many voxels (i.e., 3-D 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 (e.g., comparing neural responses for the processing of faces and of houses). Firstly, the smoothing procedure improves anatomical intersubject variability. Secondly, it increases signal-to-noise ratios by reducing the 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 3-D fixation maps generated by *i*Map shares this problem, since the 3-D fixation maps contain thousands of pixels and, therefore, also generate a large number of statistical comparisons. In addition, similarly to voxels in the fMRI space, pixels are not statistically independent; the data for a particular pixel tend to be similar to those for 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 approach is based on two main processing stages: First, it relies on the estimation of the smoothness (spatial correlation) of the statistical maps, and then these smoothness values are used to determine
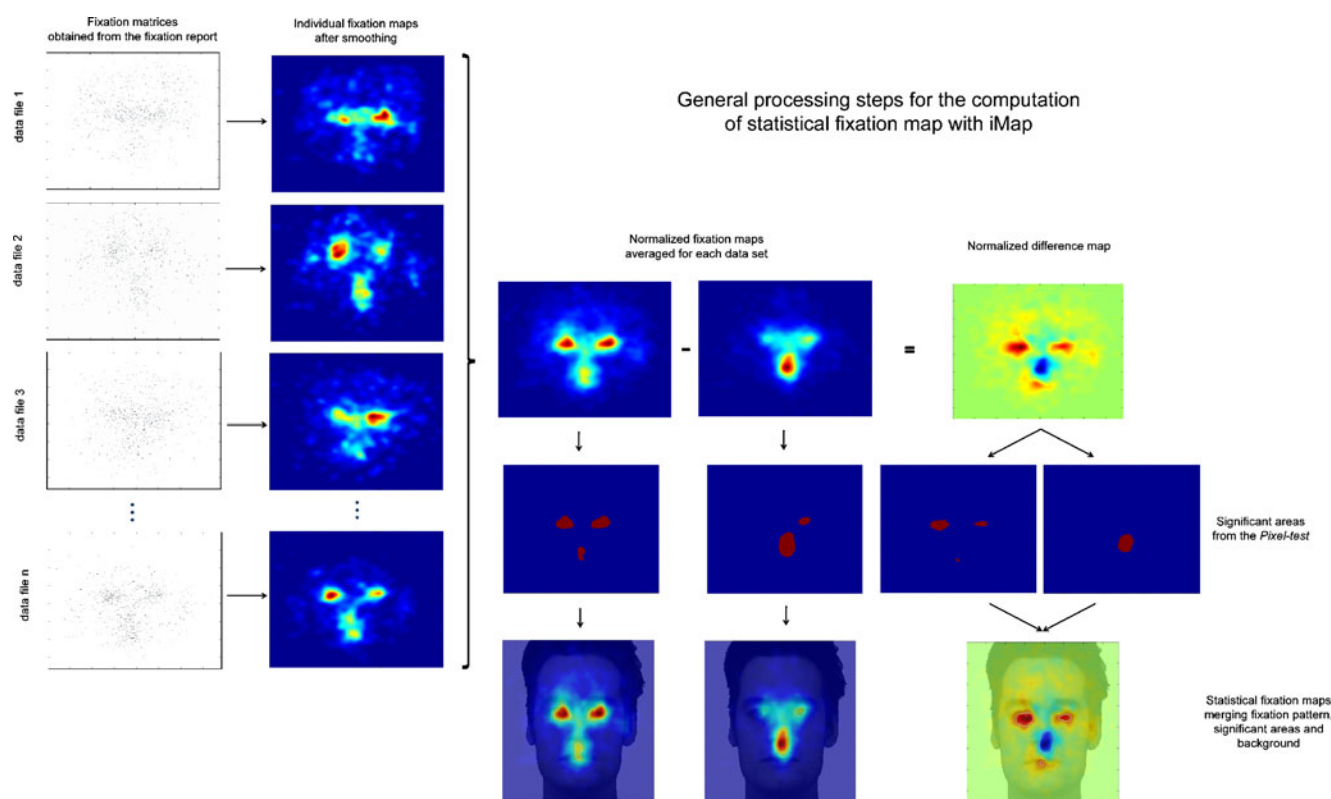
**Fig. 2** General processing steps for the computation of a statistical fixation map with *i*Map (adapted from Caldara et al., 2010). Individual fixation maps are smoothed by convolving a Gaussian kernel on each fixation. The resulting fixation maps for all trials are then averaged, resulting in a single fixation map per condition. The differential fixation map highlights significant eye movement biases. The significant areas are determined by using the Pixel test (Chauvin et al., 2005). Finally, statistical fixation maps are produced that merge the fixation patterns, the areas fixated significantly above chance level, and the background

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. *i*Map relies on spatially normalized smoothed data, which therefore satisfy the formal constraints of the RFT used in fMRI. More precisely, *i*Map applies the statistical Pixel test from the Stat4Ci toolbox (Chauvin, Worsley, Schyns, Arguin, & Gosselin, 2005), which has been developed and validated for analyzing smooth classification images. The sensitivity of the Pixel test 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 *i*Map 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 does not influence 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 of the limitations of using *i*Map, this and other caveats of the approach will be addressed in the Discussion, after we provide formal knowledge of the technique and some concrete examples.

It is worth noting that methods sharing similarities with *i*Map have been introduced by various authors before us (e.g., Barrington, Marks, Hsiao, & Cottrell, 2008; Bruce & Tsotsos, 2009; Buchan, Paré, & Munhall, 2007; Harding & Bloj, 2010; Henderson, 2003; Kita et al., 2010; Pomplun, Ritter, & Velichkovsky, 1996; Tatler, Wade, Kwan, Findlay, & Velichkovsky, 2010; Torralba, Oliva, Castelhano, & Henderson, 2006; Wooding, 2002). Here, we briefly present the characteristics that few of these methods share with *i*Map and what differentiate them from it. Tatler, Baddeley, and 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. Contrary to *i*Map, these authors did not weight the probability density functions according to 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 *i*Map, significant differences between conditions could not be localized inside the stimulus space. More recently, Tatler (2007) used Gaussian smoothing in order to generate fixation maps. Nevertheless, this author still did not

weight the density functions by the fixation duration, as implemented in *i*Map. Moreover, no statistical test was performed on individual or difference fixation maps.

Note that using a Gaussian smoothing on fixations, weighted by their durations, is not novel per se, and heat map representations have become very popular in the last few years (see, 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; Watanabe, Matsuda, Nishioka, & Namatame, 2011; Wooding, 2002). However, in most of these cases, the heat maps were used only for illustration purposes. More importantly, in comparison with *i*Map, no statistical test was performed on these fixation maps in order to locate significant effects within the stimulus space. For instance, Buchan et al. (2007) or Tatler et al. (2010) generated heat maps for visualizing eye movement patterns, but they critically relied on ROIs to compute statistics. Moreover, none of these approaches was implemented as a freely available toolbox offering the numerous statistical and descriptive analyses we provide with *i*Map. For the whole stimulus space, *i*Map computes and returns the number of fixations, the total fixation duration, the mean fixation duration, the path length, and the mean saccade length. In the areas fixated significantly above the chance level after correction for multiple comparisons, *i*Map also computes $Z$-scored fixation durations (or number of fixations) and effect sizes (Cohen's $d$). Finally, *i*Map can also produce mean fixation durations, path lengths, total fixation durations, and number of fixations separated between significant areas and the rest of the visual input.

To the best of our knowledge, the method used by Leonards et al. (2007) is the one most similar to *i*Map. These authors created fixation maps based on Gaussian kernels, generated difference maps, and used robust statistics to compare conditions. The main advantage of *i*Map over their technique is its public availability, direct access to the parameters used, and ease of use. Another difference is that in Leonards et al.'s (2007) method, each fixation is replaced by an elongated Gaussian distribution around the landing point, with a spread determined by the magnitude and angle of the saccade used to get to this location (whereas *i*Map uses a circular Gaussian). This is a very interesting way to represent the distribution of landing positions, depending of the direction and size of the saccade. However, in *i*Map, 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 an elongated) Gaussian kernel allows us to keep as much as possible to an assumption-free approach. In addition, one could argue that representing the direction of the fixation with elongated Gaussians is valid, but only

for the few milliseconds following the arrival of the saccade. Thereafter, a representation closer to the physiological constraints of the visual system, with a circular foveal projection, might be more appropriate. But, similarly to *i*Map, this approach raises novel questions: For instance, for how long is an elongated representation the most appropriate way to describe the data (for 20 ms? 40 ms?)? Is the shape of the Gaussian dependent of the task at hand, on the background information, and so forth? Finally, it is 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 *i*Map (see Fig. 7 of Nuthmann & Henderson, 2010).

Other authors have used slightly different data-driven approaches. We would particularly like to mention the Scanmatch toolbox from Cristino, Mathôt, Theeuwes, and Gilchrist (2010), the scan path similarity measure of Jarodzka, Holmqvist, and Nyström (2010), and the approach used by Mannan, Kennard, and Husain (2009). Interestingly, Cristino et al. and Jarodzka et al. used similar methods: the Needleman–Wunsch algorithm for Cristino et al., and the Levenshtein distance for Jarodzka et al. (see also Harding & Bloj, 2010, for a use of this method). A key advantage of Cristino et al.'s method is that it is provided in an implemented, freely available toolbox, allowing researchers who are not expert in programming to use their approach.

The above approaches aim at describing and comparing sequences of fixations, an analysis that is not implemented in *i*Map. However, both Cristino et al.'s (2010) and Jarodzka et al.'s (2010) techniques return a single number reflecting how similar some fixation sequences are. Consequently, contrary to *i*Map, they do not allow for visualizing and statistically testing which 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 order of fixations. This strategy is also possible with *i*Map, since individual and difference maps can be computed for single fixations or sequences of fixations in order examine the temporal characteristics of oculomotor behavior.

Carmi and Itti (2006) also implemented various metrics for comparing fixation distributions. These approaches are particularly interesting and appropriate for the question they investigated, which was quantifying the agreement between human attentional selection and attention priority maps. We will not detail these metrics here, because none of them allows, as does *i*Map, for visualizing the effects at particular locations 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.

The approaches mentioned above raise the interesting point that eye movements do not provide unequivocal evidence about the measure of visual information being used by observers(Posner, 1980). Hence, it is important to stress that *i*Map generates statistical fixation maps and not so-called *attentional maps* (refer to, e.g., www.attentiontool. com or http://eyequant.com). The Gaussian kernel is a variable that can be flexibly adjusted to characterize both the visual information that can be sampled for a given fixation and/or the potential error due to the eyetracker accuracy. Therefore, the *i*Map user can flexibly set the kernel size according to the specific hypothesis, materials, 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 eyetracker. This is the approach we adopted in our previous articles, where the kernel size was set to 10 pixels, which in our setup related to the 0.5°-of-visual-angle accuracy of our EyeLink 1000 desktop eyetracker. Importantly, the ability to manually adjust this parameter allows researchers to adapt the kernel size not just to their equipment, but to the participant as well. Hence, it is possible to set a specific kernel for each individual, depending on the eyetracker accuracy, as measured during the calibration procedure.

To sum up, despite its similarities to some previous approaches, *i*Map remains an original, complementary tool for analyzing eye movement data. The main difference between *i*Map and the methods previously cited is that *i*Map provides an implemented toolbox that allows users to share an identical implementation of this technique and, above all, to compute robust statistical analyses. *i*Map generates fixation distributions smoothed with Gaussian kernels, transforming 2-D fixation maps, uniquely based on fixation coordinate locations in $x,y$ dimensional space, into a 3-D fixation landscape, with $z$ reporting the intensity of fixations (weighted either by number of fixations or their durations; see Fig. 2).

*i*Map has already been used and successfully validated in a series of eye movement studies (Blais, Jack, Scheepers, Fiset, & Caldara, 2008; Caldara, Zhou, & Miellet, 2010; Jack, Blais, Scheepers, Schyns, & Caldara, 2009; Kelly et al., 2011; Kelly et al. 2011; Kelly, Miellet, & Caldara, 2010; Miellet, Caldara, & Schyns, 2011; Miellet, Zhou, He, Rodger, & Caldara, 2010; Rodger, Kelly, Blais, & Caldara, 2010). To illustrate the functionality and flexibility of *i*Map, we provide examples below from real eye movement data and diverse statistical comparisons (i.e., across different groups of observers, in the same observers but with different tasks, etc.). Importantly, *i*Map 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 *i*Map code

and also to add plug-ins for the toolbox in the future (e.g., a plug-in to generate 3-D fixation map movies or dynamic statistical tests over time). The relevant information and files will always be freely available to use and to download online (go to www.unifr.ch/psycho/ibmlab/). Finally, *i*Map can be used on data acquired with any eyetracker and preprocessing software that can provide a fixation report, which includes the coordinates and duration of each fixation, as well as an item number.

## Method

### Installation and credits

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

The exportfig code was written by Ben Hinkle in 2001 (bhinkle@mathworks.com) and can be downloaded from www.mathworks.com/matlabcentral/fileexchange/727.

The CiVol and HalfMax routines are part of the Stat4Ci toolbox, which allows for performing the Pixel and the Cluster tests, both based on RFT. The Stat4Ci toolbox is free and can be downloaded from www.mapageweb. umontreal.ca/gosselif/basic%20Stat4Ci%20tools/.

If you use the statistical functions of the Stat4Ci package called with *i*Map (i.e., the Pixel or Cluster tests), please cite Chauvin et al. (2005), listed below in the References.

The stat_threshold function was written by Keith Worsley for the fmristat toolbox, which is free for download at 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 add them to the MATLAB path.

The Western Caucasian face images used in the examples below of how to use the *i*Map toolbox belong to the Karolinska Directed Emotional Faces database (KDEF; Lundqvist, Flykt, & Öhman, 1998).

### Creating the input matrices

*i*Map can be used with eye movement data collected with any eyetracker. First, the data need to be preprocessed in order to determine saccades and fixations. This filtering process should be feasible for the majority of the analysis software provided with eyetrackers, or alternatively, with a saccade detection algorithm based on eye velocity.

The file resulting from this preprocessing will be used as input data for the *i*Map function. The input data are a set of matrices with a single fixation per line. The only data

Fig. 3 Validation of the fixation map area when using *i*Map
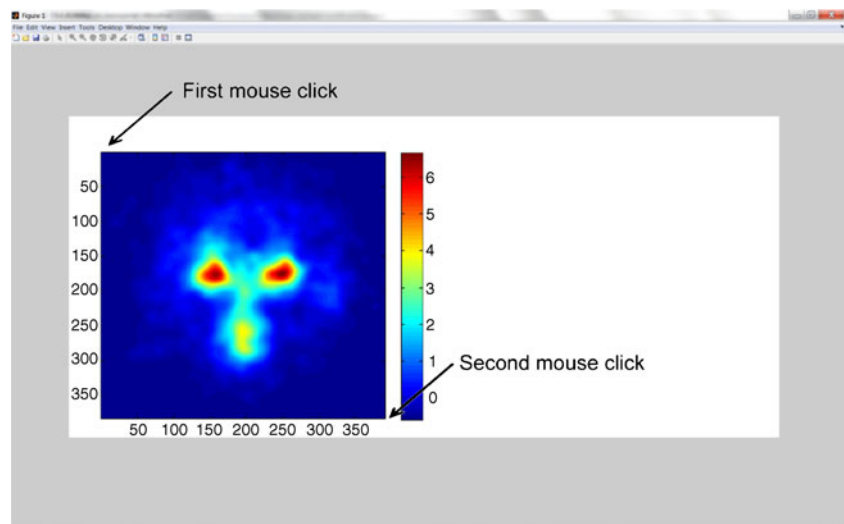


required are the coordinates and duration of the fixations and the item numbers. The order of the columns has no importance, as they have to be specified in the *i*Map function. Any other columns can be used for specifying experimental conditions. A specific input data matrix has to be created for each participant and/or condition. The input files used by the *i*Map 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 created from any .txt file (e.g., a fixation report from the EyeLink Data Viewer). Some of the examples below show how to create such input data matrices.
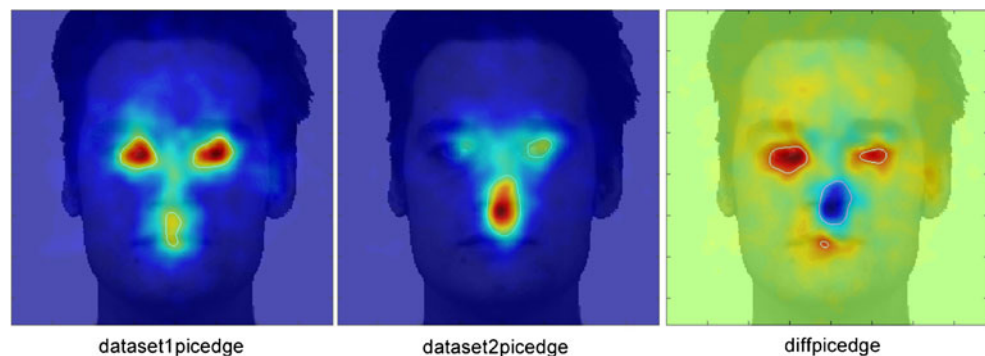
Running the *i*Map function

*i*Map 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).

The parameters of this function are defined as follows:

1. xSize, ySize: These parameters specify the stimulus size in pixels (e.g., 382, 390).

2. columnx, columny, columnduration, columnitem: These specify the column numbers for $x, y$ coordinates, fixation durations, and item number. Specifying these parameters allows for flexible data format.

3. dataset 1, dataset 2: These specify the data .mat files that will be tested/compared. For example, [1:20], [21:40] will compare Data Items 1–20 with Data Items 21–40. The second dataset is optional; this field has to be left empty if only one dataset is tested. If only one dataset is tested, *i*Map produces a statistical map and eyetracking indexes for this dataset. If two datasets are specified, *i*Map provides the statistical maps and eyetracking indexes for both datasets and the difference map and indexes.

4. standard deviation: The 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. With empty square brackets [] for this setting, the default value will be used.

5. maptype: Values are 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 centered and a central fixation cross is presented before
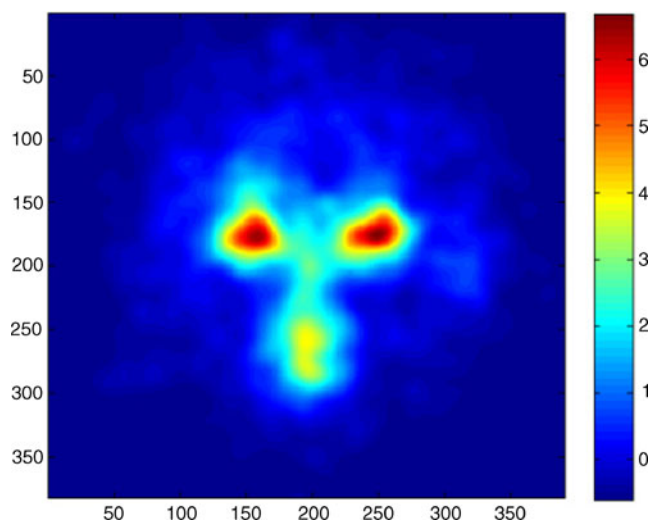
Fig. 4 Statistical fixation maps for the first and second datasets and for their difference

**Fig. 5** The map in dataset1map.tif, with scale

the trials. 1 (the default option) keeps all of the fixations, 2 ignores the first fixation of each trial.

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

8. specificfix: This setting is used to select one or several specific fixations—for example, [3 3] or [1 3]. This value is optional.

9. searchspace: By default, this is set to the size of the stimulus, xSize * ySize. The search space size can be specified by directly indicating 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 you launch the function, a map will appear on the screen. To start using the program, maximize the map, click on its top left and then bottom right corners, then hit Enter (see Fig. 3).

Output

*i*Map creates .tif pictures of the single and difference fixation maps, called dataset1picedge.tiff, dataset2picedge.

tiff, and diffpicedge.tiff, respectively. These maps can be merged with a background picture. They display the significant areas based on a Pixel test. *i*Map also creates. tif pictures of the scales of the Z-scored fixation measures, called dataset1map.tif, dataset1map.tif, and Zdiffmap.tif, respectively.

In addition, *i*Map generates .txt files with global eyetracking measures for both datasets (called eyebasicdataset1.txt and eyebasicdataset2.txt). The columns in these files are the number of fixations, the total fixation duration (in seconds), the mean fixation duration (in seconds), the path length (in pixels), and the mean saccade length (in pixels). The lines correspond to the raw data files (participants, sessions). *i*Map also creates a text file called Zscore.txt that includes the mean Z scores in the significant area for (in respective columns) Dataset 1, Dataset 2, Dataset 1 in Areas 1 and 2 (the areas in which the fixation durations are significantly longer for Datasets 1 and 2, respectively), and Dataset 2 in Areas 1 and 2.

*i*Map also produces a .txt file with Cohen's *d* values (Cohen, 1988) for both datasets in Areas 1 and 2. This file is called cohend.txt. Finally, imap creates .txt files with the eyetracking data in both the significant areas and the rest of the picture. These files are called eyeareadataset1.txt and eyeareadataset2.txt and are organized in 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 organized according to the same logic.

## Examples

For convenience, *i*Map and the supporting functions have been copied to each of the folders in these examples.

Example 1

Example 1 uses a subset of data from Caldara et al. (2010).

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.,

| | Number of Fixations | Total Fixation Duration | Mean Fixation Duration | Path Length | Mean Saccade Length |
|---|---|---|---|---|---|
| **Table 1** Output in eyebasicdataset1.txt: Averages of global eyetracking measures for Dataset 1 | 29.09677 | 7.483355 | 0.2343271 | 1,539.18 | 48.12055 |
| | 27.87097 | 5.932745 | 0.2136969 | 1,699.69 | 61.15729 |
| | 27.67742 | 7.76056 | 0.2542954 | 1,674.231 | 54.56829 |
| | 28.41935 | 7.174361 | 0.2308422 | 1,353.442 | 43.39324 |
| | 27.87097 | 4.96008 | 0.1603119 | 2,130.377 | 69.2846 |
| | 27.32258 | 7.850284 | 0.2606777 | 1,726.404 | 57.04416 |

**Table 2** Output in eyeareadataset1.txt: Averages of eyetracking measures for significant areas in Dataset 1

| Mean Fixation Duration (s) | | | Path Length (pixels) | | | Total Fixation Duration (s) | | | 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 |

1998) and the Asian Face Image Database (Bang, Kim, & Choi, 2001). Presentation was gaze contingent, with a 2°, 5°, or 8° Gaussian aperture around the fixation location. For this example, the aperture size is 8°. A central fixation cross was presented before each trial, then the 382 x 390 pixel stimulus was randomly placed on an 800 x 600 screen. Eye position was recorded every 8 ms 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°/s; saccade acceleration threshold = 9,500°/s.

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

The 1st contrast aims at comparing the respective eye movement strategies deployed by WC [2 3 8 12 13 18] versus EA [22 23 24 32 33 34] observers when learning human faces (WC and EA face stimuli for both groups). The values for maptype 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 *i*Map function could 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 Fig. 4.

This example shows the presence of significant fixation biases across the two group of observers (i.e., areas delimited by white borders). WC observers showed a fixation bias toward the eyes and mouth (dataset1picedge, red color in the diffpicedge difference map), whereas EA observers showed a fixation bias toward the center of the face (dataset2picedge, blue color in the difference map).

The scaling could be obtained from the dataset1map.tif, dataset2map.tif, and Zdiffmap.tif files. Figure 5 shows dataset1map.tif with its scale.

The numerical outputs produced by the analysis are reported in Tables 1, 2, 3 and 4. Here, for simplicity, we report only the global eyetracking measures and the measures for the significant areas in Dataset 1.

For this particular task, it was appropriate to specify only the face area as the search space (excluding the white background). Indeed, if the default search space contains a large number 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 limit the search to pixels containing information (here, a face mask):

> 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').

Reducing the search space increases the threshold, thus decreasing the sensitivity, as highlighted by the *i*Map analysis reported in Fig. 6.

Example 2

This example uses a subset of the data from Miellet et al. (2010).

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

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

| Single Maps | | Difference Map | | | |
|---|---|---|---|---|---|
| Dataset 1 | Dataset 2 | Dataset 1, Area 1 | Dataset 1, Area 2 | Dataset 1, Area 1 | Dataset 1, Area 2 |
| 4.483855 | 4.909454 | 4.646505 | 2.261218 | 1.699236 | 5.669345 |

**Table 4** Output in cohend.txt: Effect sizes in the significant areas on the difference map

| Cohen's d Area 1 | Cohen's d Area 2 |
|---|---|
| 1.575624 | −1.697319 |

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 eyetracking data were recorded in MAT-LAB, and preprocessing 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 the trials could be considered individually (please refer to the singlescenes.m Matlab code to obtain the details of this procedure) or fixation positions could be normalized relative to the target position, by creating a new fixation space where all the targets were centered in the middle of the screen (please refer to the normalizedscenes.m Matlab code to obtain the details of this procedure).
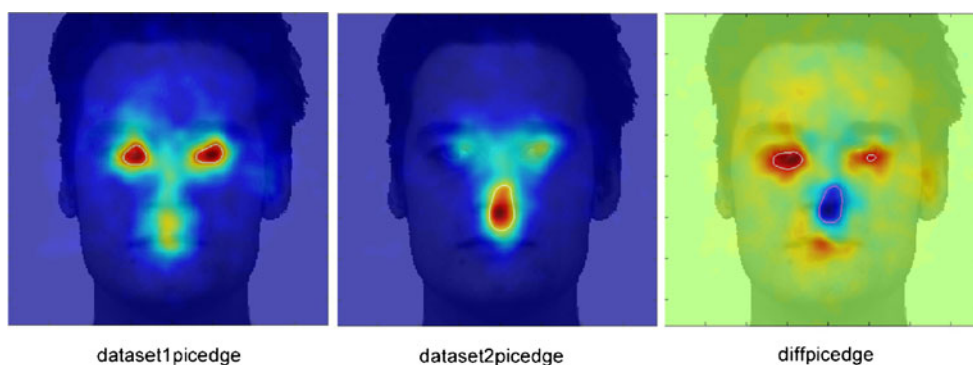
The function singlescenes.m shows how to generate data1... datan files containing the summary matrix with gaze coordinates, fixation durations, and item numbers. Here, we created such files only for the no-Blindspot (0°) and 5° target conditions. Moreover, because there is no spatial normalization in this example, we selected only the fixations corresponding to a specific item (here, Item 49). In this example, only one dataset (including 10 participants) was considered, and the first fixation of each trial was excluded (there was a central fixation cross before each trial, and the stimuli covered the full screen). Hence, the iMap function could be executed by typing the following sequence,

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

producing the scene shown in Fig. 7.

The function normalizedscenes.m is constructed similarly to singlescenes.m, aside from the fact that here there is no need to filter the data for a specific trial, as the target positions were spatially normalized. There was

also no background, as different target stimuli were considered. The iMap function could be executed by typing the following sequence,

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

producing the image seen in Fig. 8.

Figure 9 shows an example of a selection of specific fixations (6–15 of each trial, which were 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]) would also be also accepted.

This example reveals fewer surrounding fixations related to 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 an 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 and Kessler (2009).

This experiment involved learning letter positions on a 2 x 3 grid, testing high- versus low-checking participants who were presented correct versus incorrect probes (see Fig. 10). The experiment was presented with E-Prime. The raw data were recorded in SR Research's .edf format and then preprocessed using the SR Research DataViewer, with the fixation report exported in .txt format. The memorytask.m function prepares the data for analysis and runs iMap. The screen-based coordinates are also centered on the stimulus.

The preparation code (memorytask.m) allows for considering specific conditions. We can specify Probe 1 as a correct or incorrect probe (corP1 or incorP1) and several time periods (Period 1 < 2 s; 2 s < Period 2 < 4 s; 4 s < Period 3). The fixation maps reveal specific patterns for each time period.

**Fig. 6** Statistical fixation maps for the first and second datasets and for the difference map, when using a mask limiting the search space



dataset1picedge          dataset2picedge          diffpicedge

**Fig. 7** Statistical fixation map (fixation durations) for a specific item, with the first fixation of each trial excluded and the search space covering the entire stimulus/screen



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

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).

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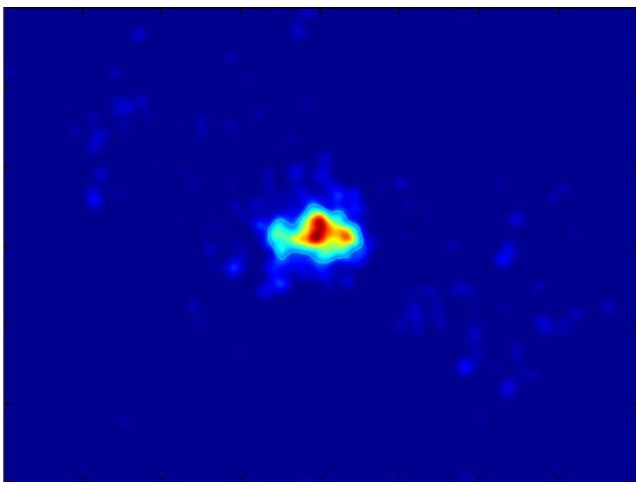
using the following parameters in the input of the *i*Map function:

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

The statistical fixation map reveals significant hotspots on each of the six 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. 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 participants.

Interestingly, the fixation maps were different for low- versus high-checkers in the third time period when the probe was incorrect (see Fig. 14). 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 was sufficient for a response from the low-checkers. In contrast, the high-checkers verified on the grid (which was empty during this period) before answering.

## Discussion

We developed an alternative method for analyzing eye movement data: *i*Map. As in 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;



**Fig. 8** Statistical fixation map for spatially normalized items

**Fig. 10** Presentation of the stimuli used in Example 3 and the time course of the trials
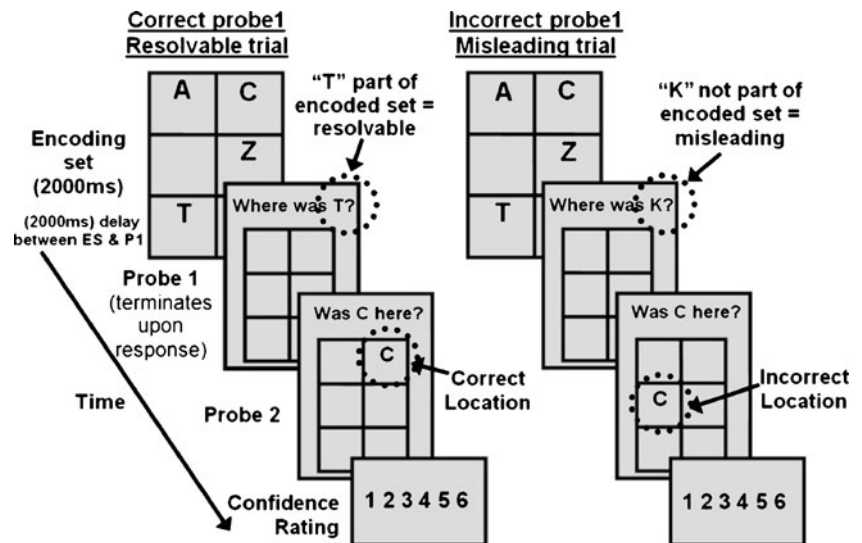


Torralba et al., 2006; Wooding, 2002), this method does not rely on the *subjective* definition of ROIs; in fact, it simply does not require the use of ROIs. Crucially, however, *i*Map offers some advantages relative to previous methods. First, it relies on robust statistics to assess the significance of effects. Second, it is coded as an editable toolbox for MATLAB, freely available for download and use (www.unifr.ch/psycho/ibmlab/).

To illustrate the functionality and flexibility of this toolbox, we have provided three examples. The results from those examples, coupled with those from our previous work on face (Blais et al., 2008; Caldara et al., 2010; Jack et al., 2009; Kelly et al., 2011; Kelly et al., 2011; Kelly et al., 2010; Miellet et al., 2011; Rodger et al., 2010) and

scene (Miellet et al., 2010) processing—which are largely consistent with the Western Caucasian (WC) eye movement literature and the literature on East Asian (EA) observers (Kita et al., 2010; Watanabe et al., 2011)—demonstrate 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 WC and 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 did EA observers, whereas EA observers fixated the central part of the face more than did WC observers. Additional analyses also revealed similar fixation patterns for both types of stimuli (WC vs. EA faces) or correct versus incorrect face
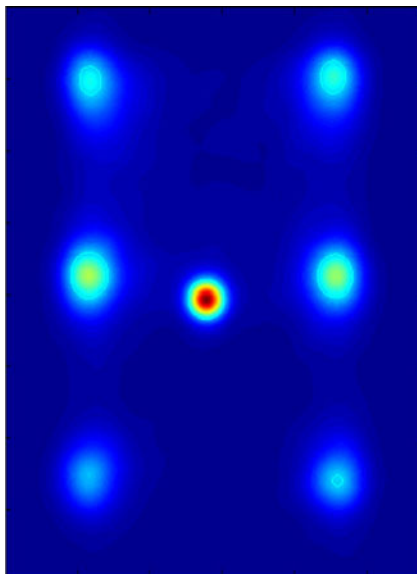


**Fig. 11** Fixation map of the learning stage, calculated across all participants and probe conditions
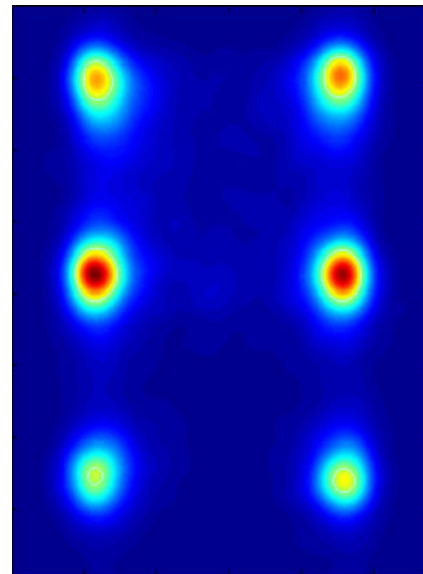


**Fig. 12** Fixation map of the learning stage, calculated across all participants and probe conditions without the first fixation
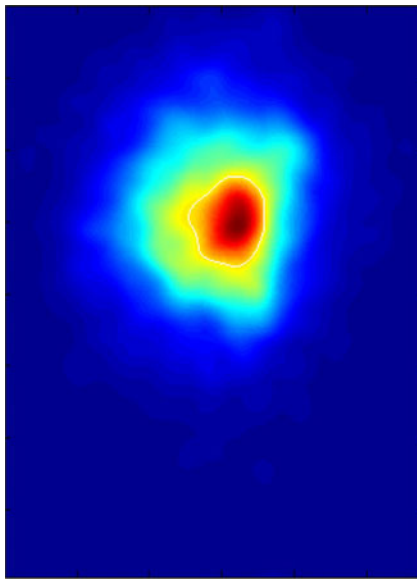
**Fig. 13** Statistical fixation map for the delay period (empty screen)

deviation of the Gaussian kernel used during the smoothing and the normalization procedures.

This parameter (i.e., width of the Gaussian kernel) is clearly dependent on the experimental stimuli used in the eye movement study and on the task at hand. In neuroimaging, this choice is perhaps easier, as the width of the Gaussian should be no larger than the brain area or region of interest; for instance, it would be between 3 and 6 mm for full width at 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 12 mm for full width at 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 not to "oversmooth" our data. We 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., 2011; Kelly et al., 2011; Kelly et al., 2010; Miellet et al., 2011; Miellet et al., 2010; Rodger et al., 2010), which is roughly the size of a fourth of the fovea (Hood & Finkelstein, 1986). We thought that this was a sensible choice for the question we aimed to address: investigating cultural diversity in face processing. However, this would not be the optimal parameter for analyzing *any* eye movement task. For instance, let's assume that we would like to investigate the role of 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 governing rule for defining the size of the standard deviation of the Gaussian kernel. Experimenters must evaluate and back up their choice with existing knowledge from the literature and also justify their choices with respect to the stimuli, the equipment, and the task used in the experiment.

recognition performance (see Caldara et al., 2010). In the second dataset, observers had to detect and identify an animal in a natural visual scene. We presented *i*Map analyses for the natural digital images and for digital images normalized (centered) on the position of the animal. These analyses showed significant fixation hotspots on the search target. Additional analyses revealed fixations on the targets despite large Blindspots and similar fixation patterns for EA versus WC participants (see Miellet 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-checking observers in a memory task. This analysis showed that high- and low-checkers deploy different strategies when confronted with ambiguous/erroneous information.

*i*Map was inspired by methods in fMRI and it suffers for very similar caveats, which we will address in turn, beginning with the choice of the width of the standard

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



Low-checkers

High-checkers

Difference map
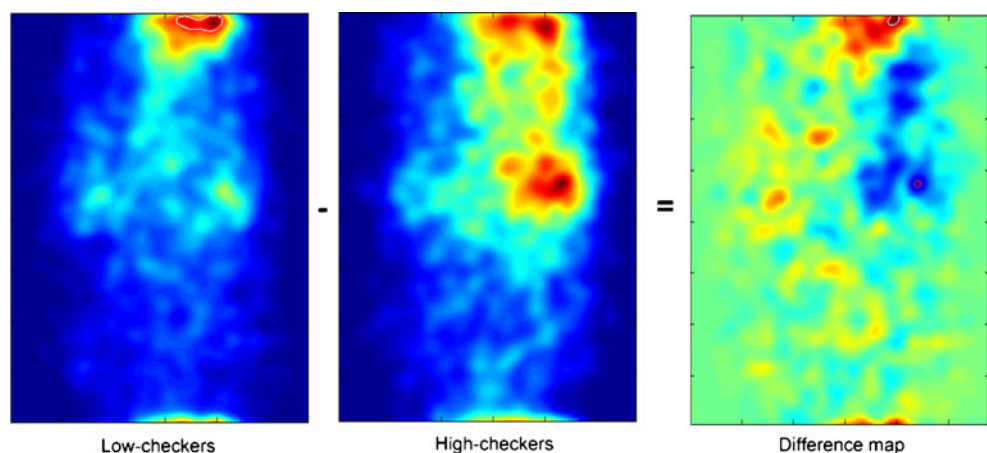
As in neuroimaging with data recorded from brains of different participants, *i*Map requires a normalized space in which to perform statistical analyses. Without entering into the details of this procedure in fMRI, there are several approaches used to normalize the human brains of different participants. The most commonly used approaches involve the realignment of brains into the Tailairach (Talairach & Tournoux, 1988) or the Montreal Neurological Institute (MNI) standard template spaces. This process ensures that the comparison of voxels across participants will be valid. Similarly, *i*Map requires that the fixation landscapes created for a particular condition be built on a homogenous space, such that a given fixation for one item is fully comparable to a fixation on the same location for 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., 2011; Kelly et al., 2011; Kelly et al., 2010; Miellet et al., 2011; Rodger et al., 2010), and in the present examples, were normalized for their eye and mouth positions. In the example we provided for visual scenes, we demonstrated that is also possible to normalize natural scenes, by arbitrarily centering the object of interest in the middle of the scene (see Miellet et al., 2010); note that the task used here was to find and identify an animal. However, *i*Map does not prevent the analysis of a unique input space, as long as many eye movement samples are collected for a particular input space to ensure the statistical validity of the analysis.

It is worth noting that *i*Map has been developed to analyze *where* and *when* eye movements are performed by observers. As illustrated by the previous examples, *i*Map 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 scan paths used by the observers (e.g., Cristino et al., 2010; Jack et al., 2009).

To sum up, *i*Map can analyze 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 *i*Map method updated, and will freely provide new versions of the MATLAB toolbox code online (www.unifr.ch/psycho/ibmlab/). We believe that various approaches and methods are necessary in any scientific discipline, so that researchers will be able to flexibly use the most appropriate method to answer the question at hand. We hope that users will help us improve *i*Map and eventually build bridges with other data-driven MATLAB-based toolboxes for eye movement analysis.

## References

Adler, R. J. (1981). *The geometry of random fields*. New York: Wiley.

Bang, S., Kim, D., & Choi, S. (2001). Asian face image database. Pohang, Korea: Pohang University of Science and Technology, Intelligent Multimedia Lab.

Barrington, L., Marks, T. K., Hsiao, J. H., & Cottrell, G. W. (2008). NIMBLE: A kernel density model of saccade-based visual memory. *Journal of Vision, 17*(14), 1–14. 8(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*, 1089–1105.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE, 3*, e3022.

Bruce, N. D., & Tsotsos, J. K. (2009). Saliency, attention, and visual search: An information theoretic approach. *Journal of Vision, 9*(5), 1–24.

Buchan, J. N., Paré, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience, 2*, 1–13.

Buswell, G. T. (1935). *How people look at pictures: A study of the psychology of perception in art*. Chicago: University of Chicago Press.

Caldara, R., & Seghier, M. L. (2009). The fusiform face area responds automatically to statistical regularities optimal for face categorization. *Human Brain Mapping, 30*, 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*, 313–319.

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the "spotlight" reveals universal information use for face recognition. *PLoS ONE, 5*, e9708. doi:10.1371/journal.pone.0009708

Carmi, R., & Itti, L. (2006). The role of memory in guiding attention during natural vision. *Journal of Vision, 4*, 898–914. 6(9).

Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M., & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision, 1*, 659–667. 5(9).

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cristino, F., Mathôt, S., Theeuwes, J., & Gilchrist, I. D. (2010). ScanMatch: A novel method for comparing fixation sequences. *Behavior Research Methods, 42*, 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. *Journal of Vision, 8*, 1–17. 10(2).

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 in Cognitive Sciences, 7*, 498–504.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition, 33*, 98–106.

Hering, E. (1879). Uber 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. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance, vol. 1: Sensory processes and perception* (pp. 5.1–5.66). 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. *Current Biology, 19*, 1543–1548.

Jarodzka, H., Holmqvist, K., & Nyström, M. (2010, March). *A vector-based, multidimensional scanpath similarity measure*. Paper presented at the Eye-Tracking Research and Applications (ETRA) conference, Austin, TX.

Javal, É. (1879). Essai sure la physiologie de la lecture. *Annales d'Oculistique, 82*, 242–253.

Kelly, D. J., Jack, R. E., Miellet, S., De Luca, E., Foreman, K., & Caldara, R. (2011). Social experience does not abolish cultural diversity in eye movements.

Kelly, D. J., Liu, S., Rodger, H., Miellet, S., Ge, L., & Caldara, R. (2011). Developing cultural differences in face processing. *Developmental Science*.

Kelly, D. J., Miellet, S., & Caldara, R. (2010). Culture shapes eye movements for visually homogeneous objects. *Frontiers in Perception Science*. doi:10.3389/fpsyg.2010.00006. 1(6).

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*, e11050.

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience, 12*, 535–540.

Lamare, M. (1892). Des mouvements des yeux dans la lecture. *Bulletin et Mémoire de la Societé Franç-aise 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. *Current Biology, 17*, R8–R9.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces [Database]*. 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. *Current Biology, 19*, R247–R248.

Miellet, S., Caldara, R., & Schyns, P. (2011). *Local Jekyll and global Hyde: The dual identity of face identification*. Psychological Science.

Miellet, S., Zhou, X., He, L., Rodger, H., & Caldara, R. (2010). Investigating cultural diversity for extrafoveal information use in visual scenes. *Journal of Vision, 21*(6), 1–18. doi:10.1167/10.6.21. 10(6).

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*, e259.

Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision, 20*, 1–12. 10(8).

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. *Journal of Neuropsychology, 2*, 245–268.

Over, E. A., Hooge, I. T., & Erkelens, C. J. (2006). A quantitative measure for the uniformity of fixation density: The Voronoi method. *Behavior Research Methods, 38*, 251–261.

Pomplun, M., Ritter, H., & Velichkovsky, B. (1996). Disambiguating complex visual information: Towards communication of personal views of a scene. *Perception, 25*, 931–948.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*, 1457–1506.

Rodger, H., Kelly, D., Blais, C., & Caldara, R. (2010). Inverting faces does not abolish cultural diversity in eye movements. *Perception, 39*, 1491–1503.

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*, 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. *Cerebral Cortex, 16*, 574–586.

Smith, S. M. (2003). *Preparing fMRI data for statistical analysis*. Oxford: 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. *Journal of Vision, 4*, 1–17. 7(14).

Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of fixation selection: Effects of scale and time. *Vision Research, 45*, 643–659.

Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2010). Yarbus, eye movements, and vision. *i-Perception, 1*, 7–27.

Torralba, A., Oliva, A., Castelhano, 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. *Psychological Review, 113*, 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*. Amsterdam: 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*, 793–804.

Watanabe, K., Matsuda, T., Nishioka, T., & Namatame, M. (2011). Eye gaze during observation of static faces in deaf people. *PLoS ONE, 6*, e16919.

Wooding, D. S. (2002). Eye movements of large populations: II. Deriving regions of interest, coverage, and similarity using fixation maps. *Behavior Research Methods, Instruments, & Computers, 34*, 518–528.

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited—again. *NeuroImage, 2*, 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. *Human Brain Mapping, 4*, 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*, 429–441.