Akori: A Tool Based in Eye-Tracking Techniques for Analyzing Web User Behaviour on a Web Site

Felipe Vera
Department of Industrial Engineering
University of Chile
Santiago, Chile
felivera@ing.uchile.cl

Víctor D. Cortés
Department of Industrial Engineering
University of Chile
Santiago, Chile
vcortes@ing.uchile.cl

Gabriel Iturra
Department of Computer Science
University of Chile
Santiago, Chile
gabrieliturrab@ug.uchile.cl

Juan D. Velásquez
Department of Industrial Engineering
University of Chile
Santiago, Chile
jvelasqu@dii.uchile.cl

Pedro Maldonado
Physiology and Biophysics Program
University of Chile
Independencia, Chile
pedro@neuro.med.uchile.cl

Andrés Couve
Physiology and Biophysics Program
University of Chile
Independencia, Chile
andres@neuro.med.uchile.cl

***Abstract*—As the use of the Internet grows every year, e-commerce's usage does as well. There is a tough competition between companies to be able to attract customers to use their services.The design of a website is crucial to retain a customer, and a retained client is more valuable over time, so understanding what attracts the attention of a potential client on a website is really important.**

This work proposes a novel web platform for understanding the most important features of a website for the user, based on biometric information provided by eye-trackers and electroencephalogram. Akori platform offers three services for understanding the most important part of a web page to the user. The first is the visual attention map, which highlights in different colors the most attractive zones for the user. The second service is a visual attention map too, but it uses a grey-scale gradient instead of colors. The third service, uses the salience map to identify the Website Key Objects on a web page and highlight the objects that are predicted as such.

Our platform is useful to the telecommunication and advertising industries, as interviews with companies managers reveal. Thus, Akori promises to be a fundamental part for planning website design.

***Index Terms*—Web Preferences, Web Usage Mining, Website Key Objects**

I. Introduction

The number of people using Internet has grown over the last years. In March 2017, Internet users were over 3 billion [1], up from 1.8 billion in 2009 [2]. In Chile, Internet penetration is about 80%, being one of the highest on the region [3]. Alongside Internet penetration, e-commerce has grown as well, only increasing year by year since 2000 up to US$25.4 billion in 2015, according to the Santiago Chamber of Commerce [4].

Consequently, companies increasingly need to have a better on-line presence, in order to take position in the market. Websites are the communication channel with customers, since they are the visible face of many companies. Companies websites must be attractive to potential customers. That is, they should be able to capture the attention of users. This same situation occurs with marketing campaigns [5]. That implies that websites must be well structured and designed for accomplish a good relation with the customer.

Advertising campaigns commonly used market research techniques, such as focus groups or surveys. In addition to those traditional tools, generic technological tools, like Google Analitycs, are used with the aim of obtain basic statistics of user behavior on a website. Both techniques for measuring customers have some disadvantages. Regarding to traditional techniques like surveys, they lack in accuracy and they are expensive. The second techniques, such as Google Analytics, they only measure the number of clicks when sites are already operational, so it a ex-post solution.

So, the importance of design is crucial to retain customers of a web site, and a retained customer is more valuable over time [6]. Thus, as e-commerce is more relevant every year, it is more important to identify the characteristics that capture web user attention, in order to construct web sites interesting for them.

With the aim of helping to the development of better web sites design, we present a web platform based on web intelligence and visual attention modeling for improving visual exploration of the structure and content of sites web. This is the Advanced Kernel for Ocular Research and Web Intelligence (Akori) platform. In this website, the customer can type its company web site, and the platform compute three different maps; two of them highlights salience zones from the website and the last one highlights website key objects from the site.

Akori platform introduces a novel approach for identifying the most relevant characteristics of a website. It is based in the information provided by biometric tools, in specific eye-tracking tools. Eye-tracking tools take the data of eye movements and pupil dilatation, for identifying the objects that capture human eye attention. That way, Akori is an accurate tool for specifying important website objects.

The rest of this document is structured as it follows. Section

2375-9259/17 $31.00 © 2017 IEEE
DOI 10.1109/ICDMW.2017.90

2 contains a brief survey of the subjects that this paper cover and different applications that exist. Section 3 explains the functionality offered by our system and the algorithms involved. Section 4 describe the logical architecture of the system and the technologies used for it. Section 5 show some test that we run with our system. Finally, section 6 presents our conclusions and ideas to improve the platform on the future.

II. Related Work

A. Biometric Information Fusion for Web User Navigation and Preferences

Understanding web user preferences and navigation habits has been a research topic over the past years. This field of investigation is usually encompassed on Web Usage Mining [7], that is defined as *"the application of machine learning techniques over Web data for automatic extraction of behavioural patterns from Web users"* [8]. In the beginning, this research field focused on different techniques to model behavior and preferences of web users. Techniques for accomplish those objectives include: the use of sources as the web log from a site, the number of sites linked to another site, the use of cookies for tracking the user navigation, the use of embedded scripts on the website, among others. In a way to obtain more objective and accurate information about web site preferences, a different approach on the last years has been surged, which includes the use of biometric features and tools, as Eye Trackers (ET) or Electroencephalograms (EEG). That way, biometric tools has been used to generate new models about the preferences from a web user, and their behavior on the navigation on a web page [9].

B. Visual Attention Modelling

In the past years, visual attention modeling has been a very active research area, in which there have been diverse types of approaches. Visual attention is a general concept, that cover all factor that influence on the attention on an image, whether they are scene-driven bottom-up (BU) or expectation-driven top-down (TD). BU is focused on the information given on the same image, like contrast, saturation, objects among others. By the other side, TD focus on the information that the person has, like the task that wants to accomplish or their educational background [10].

Applications of visual attention models are diverse, since there are many perspectives from different approaches, which use different assumptions to generate visual attention models. Despite the evolution years of research on this field, there is not a general consensus to identify which visual attention model is better than other [11]. This could be a consequence that there is not a defined language to described some topics, or that models are designed for a specific kind of problem [12]. Other causes for the lack of agreement in visual attention modeling, lies in the definitions of TD and BU approach. TD approach is subjective itself, and consequently, there is fewer investigation on that approach [10].

Regarding to BU approach, the models often uses the term *saliency* to describe some parts of an scene that seem to pull ahead over their neighboring parts. That way, *saliency* is a scene-driven factor that does not depend on the person and only on the image itself [13], [14]. Usually, BU models results reflects on saliency maps. In other words, saliency maps are simply heat maps, where the hottest areas are the more attractive ones for the observer, and the coldest ones are the areas that do not stand out among the others. Nowadays, research in saliency maps is taking flight, since there is an active ranking from MIT [15] that compares different models of visual attention with eight types of criteria. The ranking gives to researches the opportunity to choice the kind of model that could be useful.

C. Website Key Objects

In web-intelligence, according to Velasquez [16] the concept of Web Object is defined as *"a structured group of words or a multimedia file present within a web page that has metadata for describing its content"*, meanwhile a Website Key Objects are *"WebObjects or groups of WebObjects that attract the Web users attention"*. Thus, Website Key Objects represent any content from a webpage that are interesting for an user.

The identification of Website Key Object has been a long road. A first approach for identifying the elements from a web page, consisted in the application of text mining techniques, in order to obtain web site keywords. The identification of web site keywords allows the recognition of relevant content from a web page for a web user [17]. Later, a generalization of this approach was created to obtain Website Key Objects [16]. This second approach took in account not only the information that was extracted with the text, but it also considered the estimated time that one web user expended on a Web Object. On the past years, research on this topic incorporated biometric factors, like eye-trackers and electroencephalogram technologies, for identifying Website Key Objects [18]–[21] and for improving other research, like web click intention [22].

D. Web Page Visitor Visualization Systems

As the design of a web page matters for a returning visitor [6] and the use of Internet has been expanded over time [1], there has been different tools to improve the design of a site, in order to capture user attention. As far as we know, there are not recent surveys on this kind of tools; we provide some examples.

- **Eyequant**[1]: The service of this platform gives to the user a heatmap that predicts, with artificial intelligence, the fixations of the eye. This predictions are supervised by Itti and Koch, and they claim that the service predicts with about 85% of precision based on the AUC curve. The service provides an easy way to make A/B tests to check the different saliency among webpages.
- **Crazyegg**[2]: The service of this platform is be able to know where the visitor is clicking on a web page and how the site is being scrolled. With a code that the client insert

[1]http://www.eyequant.com/
[2]https://www.crazyegg.com/

on their pages, Crazyegg recollect the info and produces some heatmaps with the services that they offer.

- **SentiaLabs**[3]: This Chilean company offers personal testing on a website with different services. All their services are personalized and testing with persons, like usability test and eye tracking service. As their data is being made on every task, they offer more accurate but non instant information.
- **Clicktale**[4]: This company offers mouse centered services, that is scroll heat map, mouse track path, click intention, among other. Like Crazyegg, the service needs to install a code on every page that it wants to be tracked, and analyze the mouse usage from the visitors.

III. Proposed System

Akori platform uses a bottom-up saliency to predict what attracts the attention of a web user on a page. With this approach, the service compute three different maps that allows evaluate a web page. Each map is described as follows.

A. Visual Attention Map

This method is based on an architecture which combines features extracted at different levels of a Convolutional Neural Network (CNN), as opposed to previous works that perform a non-linear combination of features extracted from the last convolutional layer to predict saliency maps [23]–[25]. This model is composed of three main blocks: a feature extraction CNN, a feature encoding network, that weights low and high level feature maps, and a prior learning network.

This model outperforms other visual attention models with a implementation in Pyhton, and using a bottom-up approach. Though it is difficult to compare and evaluate different visual attentions models, in order to know which one is the best [12], we compare the results from the methods in the MIT Saliency Benchmark [15], where we prioritize the algorithms with a implementation on Python. The model with the higher AUC output was the Multi Level Network for saliency prediction by Cornia *et al* [26][5].

This map is the input for the next two maps.

B. Visual Focus Map

The Visual Focus Map is created from a matrix of 0 and 1, where each point is 1 if that pixel is salient for the user or 0 otherwise. To create the matrix, this map receive as input the values for the saliency that is calculated based on [26], and a threshold is defined where the value of the matrix is 1 if the same pixel of the saliency matrix is greater than the threshold, or 0 otherwise. With this matrix a new map is generated where the zones that attracts the attentions are visible without the degradation of the heat map.

[3]http://www.sentialabs.com/

[4]https://www.clicktale.com/

[5]Code is available at: https://github.com/marcellacornia/mlnet

C. Website Key Objects Map

One of the most important contributions in Identifying Website Key Objects One over the last years, is the inclusion of eye tracking and pupil dilation analysis [19]. This inclusion is helpful to determinate what segment from a web page is interesting for the user. As explained before, saliency maps are based on what capture more attraction from an image, thus, we combine the results of the saliency map and the identification of objects from a webpage to determinate where are the website key Objects [27].

As shown on the follow algorithm:

Require: *url*: url to parse
Require: X_{sal}: saliency matrix
1: *elements* $\longleftarrow$ *tags* $\in$ *url*
2: $X_{ko} \longleftarrow m \times n$ *zero matrix*
3: **for all** *element* $\in$ *elements* **do**
4: $x_1 \longleftarrow$ *element position x*
5: $x_2 \longleftarrow x_1 +$ *element width*
6: $y_1 \longleftarrow$ *element position y*
7: $y_2 \longleftarrow x_1 +$ *element height*
8: **if** $x_1 \vee x_2 \notin [0, m]$ **then**
9: continue
10: **end if**
11: **if** $y_1 \vee y_2 \notin [0, n]$ **then**
12: continue
13: **end if**
14: $Sub_{matrix} \longleftarrow X_{sal}\,[x_1 : x_2, y_1 : y : 2]$
15: **if** $Sub_{matrix} size = 0$ **then**
16: $\alpha \longleftarrow 0$
17: **else**
18: $\alpha \longleftarrow \frac{\sum x_i j \in Sub_{matrix}}{Size\ of\ Sub_{matrix}}$
19: **end if**
20: **end for**

the url and the matrix of saliency are inputs. First, the algorithm takes all the elements from a webpage inside the body tag. Then, for every element it takes their position on the page, if the position is outside the boundaries of $m \times n$ it discard that element. If the position is on the page, it takes that element and merges the object with certain opacity on a new matrix based on how many saliency that element capture form the saliency matrix.

After every element has passed, the algorithm merges the new matrix with the original image resulting on a Website Key Object Map. This map assumes that every element is rectangular, and an image on the webpage is a WebObject by itself, independent if the image has different objects on it.

IV. Architecture

The service is implemented on the Python Web Framework Django[6]. As shown on the figure 1, this framework separates the logic of the site on three components: templates, views and models. The template is the design of the pages which the users interacts, in our service that is the *index* and the page *results*. The views are the logic layer that handle the request

[6]https://www.djangoproject.com/

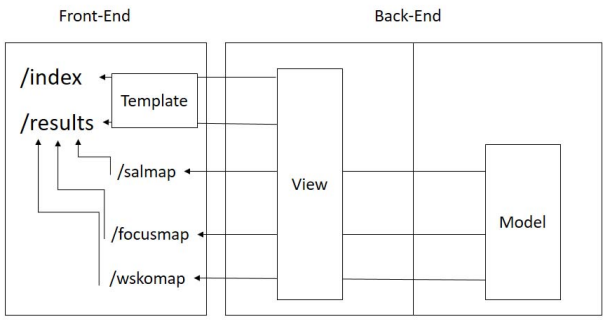

Fig. 1. A simple layout of the architecture of the service

Fig. 2. Index of the service

that the user do, in this case *index*, *results*, *salmap*, *focusmap* and *wskomap* have views to handle different request. The last three, on their views, ask to the framework to process some models and are on the component with the same name. Finally, the model is the algorithm for identifying visual attention.

When an user makes a query for a website, the service ask for the saliency map with Ajax. The view uses selenium as a headless browser, and then it visits the site. Once in the website, it takes a screen-shot for the complete site and saves it on a global variable, that represents the *driver* used with selenium, and the path of the screenshot.

Then, *salmap*, ask for the model of the saliency map and generate the matrix with the data of the screenshot.

After a successful call of *salmap*, the site makes another request with Ajax for *focusmap*. This request uses the matrix obtained on *salmap*, and creates a image for enhance the interest zone from a web user.

Once the *wskomap* is requested, it takes the *driver* saved by selenium and ask for every *html* tag that is inside the body tag. Next, *wskomap* use the matrix obtained on *salmap* and heighten the web objects that fits with the saliency map. In that way, we obtain a map for the Website Key Objects. Finally, *wskomap*, creates four maps with different colors (blue, red, orange and green) for the user to choose.

V. Evaluation

For the evaluation of the service, we made interviews with seven managers from different Chilean companies grouped on two categories:

- **Group 1:** Companies that have a commercial web site, that is highly informative about their products and services, and the main focus is the auto service by the user, like telecoms companies.
- **Group 2:** Companies with a focus on e-commerce, which is the only sales channel of their products.
- **Group 3:** Advertising agencies that make online and offline marketing.

The interview was conducted by one person, whom explained the goals of the services to the managers and shows on a tablet the index of the service as on figure 2. Next, the interviewers used the services with a website of their preference, and after about 90 seconds the results was shown as on figure 3. We let the managers to use the site and talk about their thoughts and appreciations. A compilation of the results are shown on table I. The Visual Attention map and the Focus map features were the most appreciated services; meanwhile the Website Key Objects map, was lacking of the tags of the objects highlighted. On other comments, the general consensus was that an AB Testing feature is needed on this type of service, as a manager would want to know the difference of different layout on a web site.

Interesting, advertising agencies, even without websites as their core, lead on visual attention maps. The information that the first map gives to an agency, has appealing for their potential use on marketing campaigns to know what zones attracts the attention for a viewer. To use the platform in that way, it is necessary to be able to receive as input images and not only an url.

In Chile, is common to segment the customer on age, sex and socioeconomic status, as they are differentiated on each other and companies focus on some segments of the population. For that reason, a select option on how a segment visualize a website was a regular comment on the interviews. Other suggestions were to score the cluttering on a web page, to be able to identify the differences between some words (like ”buy”, ”add to cart”, ”promotion”), to be able to identify types of website as each site has a different goal, to have a online folder where you could save the analysis, among others.

VI. Conclusions and Future Work

This work proposes a web platform that offers three services to be able to understand what part of a web page is important to a web user. The user handles the URL of the website to be analysed, then the platform gets a screenshot of the web page and performs the analysis of the image. The first analysis is the visual attention map. This service is able to show the saliency map of the web page, a map that highlights what attracts most of the users attention. The second service is a modification of the first one. In this map, we use a threshold to highlight the zone that stands out on the image, but using a greyscale gradient instead of colors. The last service, uses the saliency map to identify the Website Key Objects [28] on a web page and highlight the objects that are predicted as such.

The added value that this platform gives to its potential users is evaluated and confirmed with the different surveys

TABLE I
EVALUATION OF THE SERVICE

Group	Comments on the service	Comments on further services
Group 1	The Visual Attention map was appreciated and marked as useful with the focus map. As the WSKO map, there was a sense of not giving new information.	There will be appreciated if the services had a dedicated segment to do AB Testing and a segmentation of the visual prediction.
Group 2	The Visual Attention Map and the Focus Map was useful for the managers and the WSKO map would be more useful if could give the associated html tags for that Website Key Object.	There will be appreciated if the services had a dedicated segment to do AB Testing and a selection input to choose what segment of people would you test.
Group 3	As they focus more on marketing campaigns, visual attention maps was very useful but a way to use an image as input is needed.	Some way to save the analysis is very important to have, as well an AB testing to compare different campaigns.

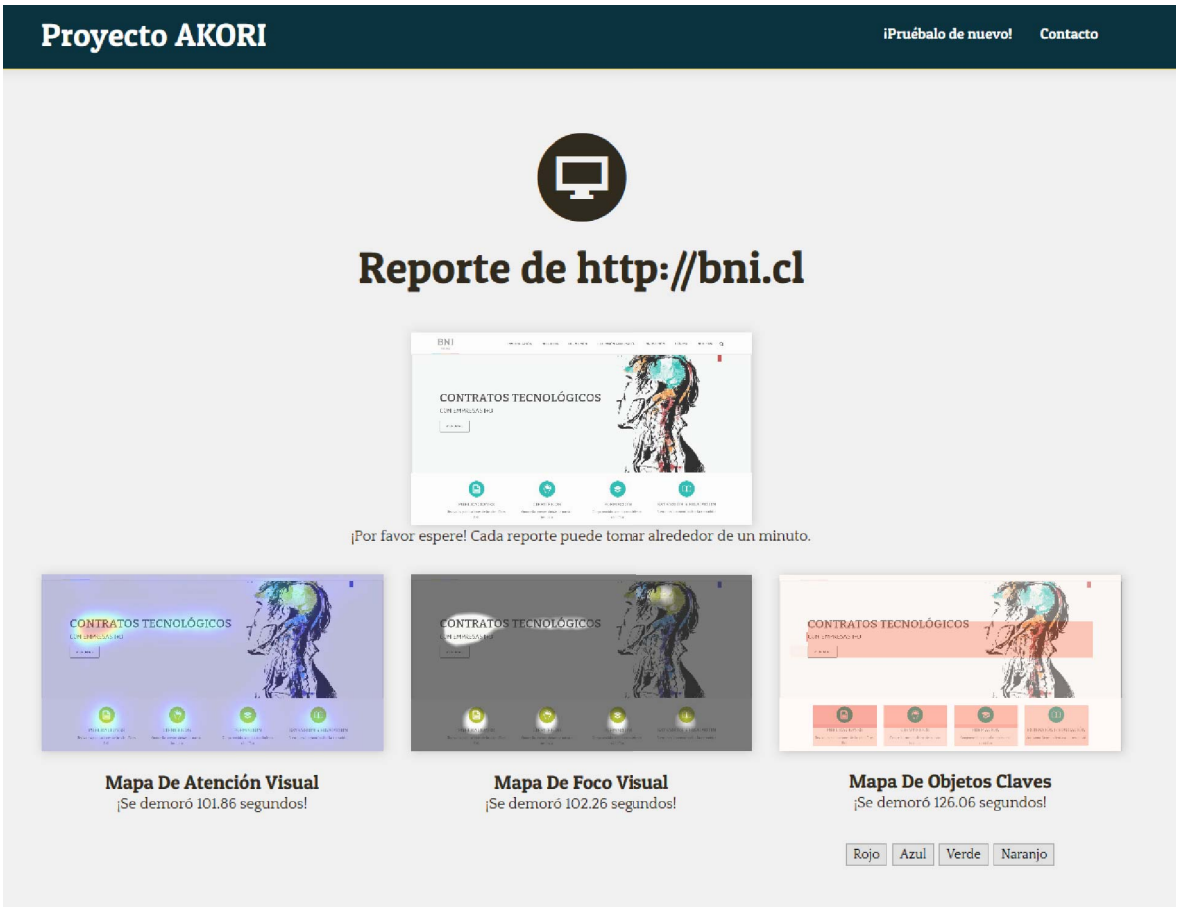

Fig. 3. Results of the service

that were carried out. Knowing what attracts the attention of a web user on a web page could be a key factor of the design of web sites or advertising campaigns. However, the proposed platform lacks some features that are needed to launch the service to the market. In the near future, it is imperative that an A/B testing comparison tool is available for customers to use and to take advantage of the full potential of a tool that delivers results in minutes; saving the *xpath* of a Web Object is necessary for a tool that identifies Website Key Objects; and last but not least, with further research on TD approach for saliency maps, different map outcomes could be achieved considering the gender or age of the user. Those features would be key to determine possible customers, as they are commonly used to define the design of a website and that kind of segmentation is widely used on marketing in Chile.

VII. ACKNOWLEDGEMENTS

This work was supported partially by Fondecyt 1160117, Fondef IT15I10021 and the Millennium Institute on Complex Engineering Systems (ICM: P-05-004-F, CONICYT: FBO16).

REFERENCES

[1] M. M. Group, "World Internet Users Statistics and 2017 World Population Stats," 2017.

[2] I. T. Union, "ICT Facts and Figures 2016," 2017.

[3] M. M. Group, "South America Internet Usage Stats, Population Statistics and Telecommunications Reports," 2017.

[4] C. d. E. d. l. E. D. Cmara de Comercio de Santiago, "La Economía Digital en Chile 2016," tech. rep., Cmara de Comercio de Santiago, 2017.

[5] C. A. Bulley, V. Adu-Brobbey, and E. O. Duodu, "Neuromarketing and the potential application of scientific methods in measuring consumer behaviour," in *Handbook of Research on Consumerism and Buying Behavior in Developing Nations*, pp. 263–282, IGI Global, 2016.

[6] D. Cyr, "Return visits: a review of how web site design can engender visitor loyalty," *Journal of Information Technology*, vol. 29, no. 1, pp. 1–26, 2014.

[7] J. D. Velasquez, A. Bassi, H. Yasuda, and T. Aoki, "Mining web data to create online navigation recommendations," in *Data Mining, 2004. ICDM'04. Fourth IEEE International Conference on*, pp. 551–554, IEEE, 2004.

[8] G. Castellano, A. M. Fanelli, and M. A. Torsello, "Web usage mining: Discovering usage patterns for web applications," in *Advanced Techniques in Web Intelligence-2* (J. D. Velsquez, V. Palade, and L. C. Jain, eds.), no. 452 in Studies in Computational Intelligence, pp. 75–104, Springer Berlin Heidelberg, 2013. DOI: 10.1007/978-3-642-33326-2_4.

[9] G. Slanzi, G. Pizarro, and J. D. Velsquez, "Biometric information fusion for web user navigation and preferences analysis: An overview," *Information Fusion*, vol. 38, pp. 12–21, 2017.

[10] A. Borji and L. Itti, "State-of-the-Art in Visual Attention Modeling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, pp. 185–207, Jan. 2013.

[11] N. Wilming, T. Betz, T. C. Kietzmann, and P. Knig, "Measures and limits of models of fixation selection," *PLOS ONE*, vol. 6, no. 9, p. e24038, 2011.

[12] Z. Bylinskii, T. Judd, A. Oliva, A. Torralba, and F. Durand, "What do different evaluation metrics tell us about saliency models?," *arXiv:1604.03605 [cs]*, Apr. 2016. arXiv: 1604.03605.

[13] C. Koch and S. Ullman, "Shifts in selective visual attention: Towards the underlying neural circuitry," in *Matters of Intelligence* (L. M. Vaina, ed.), no. 188 in Synthese Library, pp. 115–141, Springer Netherlands, 1987. DOI: 10.1007/978-94-009-3833-5_5.

[14] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.

[15] Z. Bylinskii, T. Judd, A. Borji, L. Itti, F. Durand, A. Oliva, and A. Torralba, "MIT Saliency Benchmark," 2015.

[16] J. D. Velsquez, L. E. Dujovne, and G. LHuillier, "Extracting significant Website Key Objects: A Semantic Web mining approach," *Engineering Applications of Artificial Intelligence*, vol. 24, pp. 1532–1541, Dec. 2011.

[17] J. D. Velsquez, S. A. Ros, A. Bassi, H. Yasuda, and T. Aoki, "Towards the identification of keywords in the web site text content: A methodological approach," *International Journal of Web Information Systems*, vol. 1, pp. 53–57, Feb. 2005.

[18] G. Slanzi, C. Aracena, and J. D. Velsquez, "Eye tracking and EEG features for salient web object identification," in *Brain Informatics and Health*, pp. 3–12, Springer International Publishing, 2015. DOI: 10.1007/978-3-319-23344-4_1.

[19] P. Loyola, G. Martinez, K. Muoz, J. D. Velsquez, P. Maldonado, and A. Couve, "Combining eye tracking and pupillary dilation analysis to identify Website Key Objects," *Neurocomputing*, vol. 168, pp. 179–189, Nov. 2015.

[20] J. D. Velasquez Silva, "Improvement of a Methodology for Website Keyobject Identification Through the Application of Eye-Tracking Technologies," in *Proceedings of the The 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology - Volume 03*, WI-IAT '12, (Washington, DC, USA), pp. 59–63, IEEE Computer Society, 2012.

[21] J. D. Velsquez, "Combining eye-tracking technologies with web usage mining for identifying Website Keyobjects," *Engineering Applications of Artificial Intelligence*, vol. 26, pp. 1469–1478, May 2013.

[22] J. Jadue, G. Slanzi, L. Salas, and J. D. Velsquez, "Web user click intention prediction by using pupil dilation analysis," in *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, vol. 1, pp. 433–436, IEEE, 2015.

[23] N. Riche, M. Duvinage, M. Mancas, B. Gosselin, and T. Dutoit, "Saliency and human fixations: State-of-the-art and study of comparison metrics," in *2013 IEEE International Conference on Computer Vision*, pp. 1153–1160, 2013.

[24] X. Huang, C. Shen, X. Boix, and Q. Zhao, "SALICON: Reducing the semantic gap in saliency prediction by adapting deep neural networks," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 262–270, 2015.

[25] E. Vig, M. Dorr, and D. Cox, "Large-scale optimization of hierarchical features for saliency prediction in natural images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2798–2805, IEEE, 2014.

[26] M. Cornia, L. Baraldi, G. Serra, and R. Cucchiara, "A Deep Multi-Level Network for Saliency Prediction," *arXiv:1609.01064 [cs]*, Sept. 2016. arXiv: 1609.01064.

[27] R. Rosenholtz, Y. Li, and L. Nakano, "Measuring visual clutter," *Journal of Vision*, vol. 7, pp. 17–17, Jan. 2007.

[28] L. E. Dujovne and J. D. Velásquez, "Design and implementation of a methodology for identifying website keyobjects," in *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pp. 301–308, Springer, Berlin, Heidelberg, 2009.