I. Objectives

The presentation of information can take myriad visual forms, such as bar charts, scatter plots, network diagrams, and tables. These information graphics are attempts to map potentially large amounts of complex data to easily navigable visual form for rapid and accurate knowledge transfer. However, there is not yet a satisfactory formal methodology for selecting the most appropriate visualization method for a given set of data. The research program outlined in this proposal has three primary objectives:

Objective 1: Characterize how the choice of information visualization method impacts task performance and oculomotor behavior.

Objective 2: Determine if this behavior is discriminable between experts and novices.

Objective 3: Construct an empirically-based taxonomy of visualization from *1* and *2*.

Characterizing the effects that different visualization methods have on oculomotor behavior and task performance will help lead to a better understanding of how humans look at data, respond to the relationship between data structures and visual composition, and respond differentially to visualizations of different types. The information visualization methods will be differentiated by the categories of a visual taxonomy, and the data underlying each information graphic will be characterized by a data taxonomy. Given these taxonomical characterizations of an information graphic, if we correlate eye movements (implicit skill) with problem-solving ability (explicit skill), then the first objective raises four hypotheses and the second objective raises a fifth hypothesis:

Hypothesis 1: The effectiveness of an information visualization method (based on task performance) is a function of the task.

Hypothesis 2: The effectiveness of an information visualization method (based on task performance) is a function of underlying data features.

Hypothesis 3: Task performance is discriminable in accordance with the categories of a visual taxonomy.

Hypothesis 4: Eye movements are discriminable (*i.e.*, there are "signatures") in accordance with the categories of a visual taxonomy.

Hypothesis 5: Eye movements are discriminable between experts and novices.

Hypothesis 1 posits that the ability to perform a task related to an information visualization method depends on the task being performed. In other words, a given task is best performed on certain visualization types, and conversely, a visualization method is amenable to performing particular tasks. This research intends to establish these visual-task relationships for a range of visual and task categories.

Hypothesis 2 posits that the ability to perform a task related to an information visualization method also depends on the data structure (set of data features) underlying the visualization. In other words, given data are best treated by certain visualization methods, and conversely, a

visualization method is best suited to represent particular data structures. This research intends to establish these data-visual relationships for a range of data and visual categories.

Hypotheses 3 and 4 are intended to substantiate and modify the visual taxonomy, and address the third objective above. They maintain that there is a mapping from subject task performance and eye path profiles for a given task set performed on given visualizations to the visual taxonomy characterizing these visualizations. This would result in a psychologically tested taxonomy, the first known to the author. Empirically defining a visual taxonomy will benefit education and research interested in the structure of visual information, and will inform design choices made in information visualization. One implication of this research is a determination of how effective different visualization methods are at conveying information; this understanding will be of profound help to anyone interested in communicating effectively in a graphical form.

Hypothesis 5 asserts that the eye movements are not only discriminable between visualizations, but between subject skill levels.

Figure 1 outlines the experimental approach, highlighting a particular data structure and visualization method under examination. The proposed research is interested in comparing tasks (hypothesis 1), data structures (hypothesis 2), visualization methods (hypotheses 3 & 4), and subject responses (hypothesis 5).

Section II will outline the behavioral data to be acquired and the methods to be employed in their acquisition and analysis. Section III will outline the visual taxonomy used to distinguish between information graphics stimuli, the basis from which a new visual taxonomy will be empirically derived.

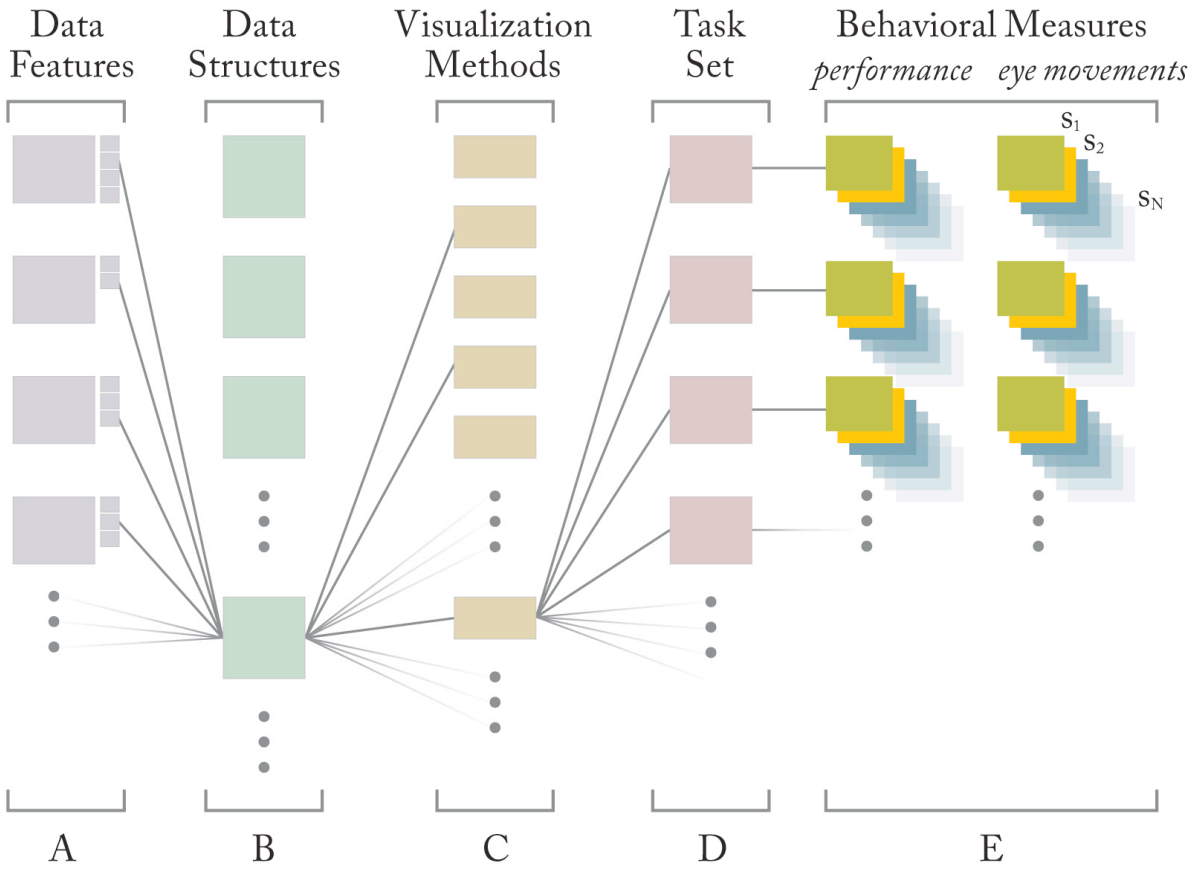


Figure 1

Methods pipeline

(A) Each data feature (scale, dimensionality, *etc.*, defined by the data taxonomy), has attributes (*e.g.*, scale may be set to nominal, ordinal, or ratio). The combination of attribute settings form (B) a data structure, which in turn is amenable to certain (C) visualization methods (defined by the visual taxonomy). When a visualization method is used to perform (D) a set of tasks, (E) performance and eye movement data are recorded for each of N subjects.

II. Methods

Stimulus acquisition and generation

Two sets of stimuli will be used for presentation: existing and generated information graphics. The existing images will be selected from a database under development by the PI whose images are being gathered from numerous image repositories on the internet, and will be classified according to the visual taxonomy in section III. The present database includes all of the images compiled by the Institute for Social Research at York University (http://www.math.yorku.ca/SCS/Gallery/), in the Atlas of Cyberspaces by the Center for Advanced Spatial Analysis at the University College London (http://www.cybergeography.org/atlas/), in Cartogram Central hosted by the National Center for Geographic Information and Analysis at UC Santa Barbara (http://www.ncgia.ucsb.edu/projects/Cartogram_Central/index.html), in the network image collection of visualcomplexity.com, as well as many other collections. Part of this research program will expand this effort to establish the most comprehensive, manually annotated (and taxonomically classified) information graphics database in the world, for use by the public via a web interface. We anticipate that it will be a valuable resource for research and for education at all levels, and that it will also be of general interest to a wide audience on the internet. In particular, the images will be important for procuring stimuli for various perceptual and cognitive psychology experiments, and will serve as a rich resource for teaching about the history and scope of visualization methods and design within and across disciplines.

The PI is in communication with collaborator Dr. Michelle Zhou of IBM to discuss the possibility of incorporating the framework underlying IBM's graphical database [18] into this research program. The advantage of doing so would be to use IBM's extremely careful manual annotations of and precise data-visual mapping between the data interrelationships and visual hierarchy for each information graphic.

The second set of images will be generated to provide flexibility of as well as control over the presentation stimuli. Flexibility refers to the requirement that from each dataset a variety of images can be constructed, according to categories of the visual taxonomy. Control refers to the requirement that the stimuli be normalized by various psychophysical criteria so as not to bias a viewer. Each stimulus will be constructed as an example of a single, predominating taxonomical category, or as a hybrid of multiple categories. The data features underlying the graphics will be based on Zhou's extensive data taxonomy [17-18]. These features include the following: scale, unit, continuity, resolution, volume, cardinality, arity, etc. From each selected combination of data features, multiple information graphics will be generated either manually (by designers) or semi-automatically (by in-house or third-party computer code).

The PI is in communication with collaborator Dr. Zhou to also discuss the possibility of adapting their computer code used to generate visual stimuli [17]. Alternatively, these stimuli will be designed by PIIM staff or similar code will be written by the PI, who has experience in programming graph-generation applications (in Matlab and in Python).

Stimulus presentation and behavioral measures

Hundreds of information graphics will be presented to each of at least 50 naïve subjects between 18 and 45 years old with normal or corrected-to-normal vision on a 21-inch color video monitor (CRT) using run-time, open source computer programs written with the aid of the Psychtoolbox (a Matlab toolbox that weds Matlab software to Macintosh hardware for precise and uninterrupted timing in perceptual cognitive experiments). During stimulus presentation, two types of subject responses will be recorded: oculomotor responses (eye tracking data) for the existing and generated images and task-based responses (mouse clicks or keyboard input) for the generated images.

The tasks will consist of answering questions about the images [1-3] involving estimation, comparison or matching between visual elements, and visual search to identify, count, or navigate through features. Half of the stimuli will be accompanied by questions posed before the trials, and half after the trials, to determine the effectiveness of a graphic in terms of how navigable it is in the former case and how memorable it is in the latter case. The questions will be of either a local or global character, to see what effects the distribution of intended fixation points in a visual search, for example, has on eye movements and task performance. Problem-solving questions used for comparing the effectiveness of a network diagram and a matrix representation of the same data might include: estimate the number of nodes and links, find a particular node or link (such as the most connected node or a link with a particular label), and find a path or common neighbor between two nodes [3].

All behavioral data will be acquired at collaborator Dr. Vincent Ferrera's laboratory at Columbia University, which conducts eye tracking and psychophysics research. All aspects of the experimental design will be made available on the PI's website as well as described in publications and presented in lectures and at conferences.

Perhaps in a later stage of this work, subjects would answer questions about the relationships between image (fixation) points to see if these overt characterizations of images agree with the classifications of the taxonomy. An extension of this idea would be to create a web interface by which a large, distributed group of people could characterize these compositional relationships as part of a game [23]. This would provide a dynamic classification system that would subsume a great diversity of perceptions and interpretations, and would accelerate classification of image databases. A subset of these inputs could be used to establish a standard database for testing purposes and a standard protocol for compositional classification of images. The range of image types may also be expanded beyond information graphics if the behavioral measures are changed, such as in recent work in spatial perception and visualization comprehension, where object properties are identified in realistic images [20].

Analysis

The eye tracking data will be analyzed by ISCAN software, and eye movement clustering will be performed using a mean shift procedure [21]. Part of this research will be to develop novel means of characterizing and comparing eye movement patterns.

If we represent data-visual mappings as graphs as in [18], another novel and potentially valuable outcome of this research will be to observe how subjects' eye movements across different visual features could be studied as a dynamic traversal through these graphs. For this, network analysis techniques would be employed to study the interaction between where subjects' eyes look and linger, and the visual composition of what they're looking at, and the structure of the data underlying what they're seeing. This network analysis is reserved for a later phase of this research.

The task-based responses will be evaluated for speed and accuracy to establish an expert group and a novice group for each task and for each information visualization method. Attempts will be made to correlate eye movement "signatures" with "expertise," not in the sense of theoretically optimal eye movements [1], but with respect to different subject eye path profiles.

One application of this expert vs. novice profiling is to automatically guide a novice to visually navigate an image in a similar manner as an expert, by carefully situating surprising elements [22] or subduing elements in the image. This is proposed as a later phase to the present research project.

III. A taxonomy of images based on composition of graphical elements

There have been numerous attempts at classifying images [4-5], based on data types and tasks [6-11], visualization algorithms and models [11-12], visual elements [13], mappings between data and visual elements [14-18], as well as subject sorting of conventional methods of presentation [19]. This section outlines a new taxonomy of images that distinguishes itself from the above approaches by detailing a formal method for classifying images based on the composition of visual elements by relational operators. The proposed bottom-up taxonomy and corresponding symbolic notation are meant to complement top-down taxonomies that classify according to content, function, intent, interpretation, or effect. The derivation of this taxonomy draws from concepts in basic geometry and linear algebra as well as graph theory and computer graphics, and draws inspiration from Gestalt psychology. As in the rest of this proposal, "image" refers to information graphic, to ensure that the relationship between image components and underlying data are well characterized. This taxonomy will act as a preliminary taxonomy; the results of this research program will be used to revise the categories of this taxonomy to come up with an empirically-based visual taxonomy.

We will consider any two-dimensional image to be a set of graphical objects set in a space and spatially related to each other. We will build up our classification system from a set of graphical elements making up these objects, the linear transforms that situate them in a space, and relational operators for relating these objects to one another in this space.

Our graphical elements will simply consist of **shapes**, which include curves[1] and boundaries[2], and **surfaces**, the simple visual qualities of a surface such as colors and textures[3]. We will consider graphical objects to consist of shapes enclosing surfaces.

Our graphical transforms are based on the affine transformations[4] used in geometry and linear algebra, where any deformation to a shape is allowed so long as parallel lines remain parallel. Example transformations include translation and changes in scale and direction (rotation, reflection, and shear). Rather than consider the transformations themselves, we will consider the effects they have on graphical objects: **position**, **size**, and **direction**.

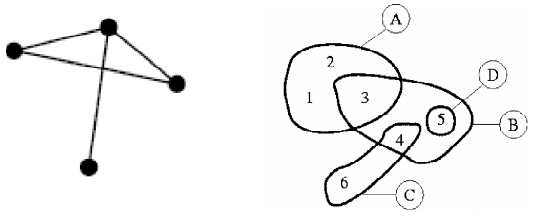
Finally, our relational operators are constructed from and act on the above elements and transforms (together referred to as attributes) or on each other, and will be of five general classes that we will refer to as **connection**, **containment**, **composition**, **correspondence**, and **context**. We will describe each relational operator below.

---

[1] A curve is a one-dimensional continuous set of points such as a line segment or the outline of a shape.
[2] A boundary is any visual property that implicitly marks the limit of an area or volume.
[3] We will refer to textures in the manner of texture mapping and raytracing in computer graphics. Virtual light sources interact with surfaces exhibiting given optical properties. These surfaces are projected to two dimensions for viewing from a given perspective.
[4] An affine transformation between two vector spaces consists of a linear transformation followed by a translation. Our transforms will sometimes deviate from affine transforms, for example when objects are scaled linearly with respect to basis vectors different than those of the image such as for radial plots. Another case is when the intent is to imply scale changes between objects, as in a caricature.

A graph[5] provides perhaps the simplest way of relating two otherwise dissimilar objects in a diagram, by simply connecting them with an edge (Figure 2). A hypergraph is a generalization of a graph, where instead of each edge connecting two vertices, each hyperedge can connect any number of vertices. The hyperedge is schematically represented in diagrams as an enclosure surrounding a subset of vertices (Figure 2). Since the taxonomy deals with visual images, we will distinguish between the visual analogs of graphs and hypergraphs to derive two different graphical relational operators: **connection** and **containment**. Connection operators are commonly seen in network and relational diagrams, where the edge of a graph can explicitly connect graphical objects, or where edges are replaced with implicit "pointers" from one graphical object to another (markers, text, etc.). Containment operators indicate the inclusion of subsets of graphical objects by the explicit use of an outline or the implicit use of boundaries (such as changing textures or surrounding shapes). Tables and Venn diagrams are examples of visuals that rely on containment to organize graphical objects.

Figure 2



simple graph *hypergraph*

Composition refers literally to the case when one graphical element or operator is made up of other elements or operators, such as a texture made up of shapes. Mosaics are canonical examples of composition of graphical objects by other objects.

Correspondence refers instead to the case where a graphical property (element, transform, or relational operator) is a function of another property. An example is when a color contour plot uses color to indicate height along an axis perpendicular to the image plane; in the taxonomical nomenclature, color is a function of position. Functional relationships refer to either explicitly defined mathematical functions or visually apparent patterns or relationships for which no apparent function exists. Correspondence is usually defined in a key or legend.

Context means the space in which graphical objects are situated, and may be founded on a graphical element, a transform, and/or an operator. For example, an element such as color could define a space such as the C.I.E. color space, the position transform characterizes the Cartesian coordinate space, and the containment operator characterizes tabular space.

[5] A graph is a mathematical construct formally relating different objects, and is depicted by a set of vertices connected by edges.

Symbolic notation

We will symbolically represent taxonomic classifications using the notation in Tables 1 and 2. Each of the symbols representing attributes of graphical objects in Table 1 are invoked in one of two cases: when there is a single dominating attribute, or when variations in an element or transform appear to be a deliberate attempt to convey meaningful information. Examples of the latter include differences in shape, color, proximity, size, or orientation. One defining feature of the notation is that a symbol that represents a varying property of an object can stand in for the object. For example, ⊕ by itself signifies graphical objects that vary in size, such as size markers in a scatter plot. Symbols for the relational operators are in Table 2; their full notational forms give rise to functional descriptions (in a mathematical or computer programming sense), where one can hierarchically embed attribute or operator symbols within any of the variables. In other words, the symbols are used as building blocks, and may also stand for groups of symbols for the purposes of compact representation and abstraction.

Table 1

Attributes of Graphical Objects

Elements		**Transforms**	
Shape	△	Position	⊙
Surface	●	Size	⊕
Shape + Surface	▲	Direction	↺

Table 2

Relational Operators*

	simple notation		full notation	
Connection	**|x|**	*"connected x"*	**|x: y; [z]|**	*"x connected by y links (links made up of z)"*
Containment	**{x}**	*"contained x"*	**w{x: y; [z]}**	*"w contains x, x enclosed by y boundary (boundary made up of z)"*
Composition	**y[x]**	*"y composed of x"*	**w[x: y; [z]]**	*"w made up of x, x enclosed by y boundary (boundary made up of z)"*
Correspondence	**x~y**	*"x a function of y"*	**x~y**	*"x a function of y"*
Context	**x/y**	*"x in y space"*	**x/y**	*"x in y space"*

* w, x, y, and z are all optional and refer to any attribute or operator

For example, a collection of right-facing arrows may collectively take the shape of a left-facing arrow or create a texture or provide a link between nodes in a network diagram, and so on. This substitution of one taxonomical category for another also allows for implicit forms of each, such as a texture boundary standing in for a curve. Taken further, any category may serve a symbolic role as well. For example, the color blue may be associated with "sky" or "water," an upper right orientation may be associated with "northeast," and a link connecting two pictures may create the impression of causal relation between the contents of the two. Because of the unlimited potential for applying semiotics to each of these categories, we will disregard the symbolic aspects of images for the purposes of classification. Taking shapes at face value creates one obvious dilemma, best exemplified by a treatment of text as graphical objects. Text functions to convey meaning apart from its form much as any visual is a symbol to be interpreted. Hereafter, text annotation is assumed to accompany almost every image and will not be incorporated in the taxonomy except in two cases: text as identifiers or as pointers, such as text labels identifying locations on a map and text pointing to another place in the image. In cases where text is the primary object, it may be denoted by "T" and where numbers are the primary objects, they may be denoted by "#," such as in a numerical table.

With regard to syntax, transform symbols usually precede element symbols and attribute symbols are sequentially ordered in a manner that conveys their successive application (such as ⊙⊕ for “positioned sizes” representing a Venn diagram of overlapping areas). Where symbols need to be clearly grouped together they are enclosed by parentheses; where they function independently they are separated by commas (such as △,● to indicate that shapes and textures convey independent pieces of information, as in a scatter plot). Additionally, numerical superscripts are applied to disparate symbols in a notational string when they need to be identified as related or identical for clarity.

To summarize, the taxonomy has five general classes of relational operators (connection, containment, composition, correspondence, and context) that act on each other or on the following five properties of graphical objects: the elements shape and surface, and the transforms position, size, and direction. The conciseness and power of this taxonomy become apparent when applying combinations of the above properties to complex data visualizations. Distinctive, canonical forms of visual presentation are neatly distinguished from one another by these bottom-up classifications and are succinctly represented by the symbolic notation of Tables 1 and 2.

[1] Foster, M.E. 2003. Evaluating models of visual comprehension. *Proc. of Eurocogsci '03; The European Cognitive Science Conference 2003*. 115 (6).

[2] Peebles, D., Cheng, P.C.-H. 2003. Modeling the effect of task and graphical representation on response latency in a graph reading task. *Human Factors.* 45(1): 28-45.

[3] Ghoniem, M., Fekete, J.-D., Castagliola, P. 2004. A comparison of the readability of graphs using node-link and matrix-based representations. Proc. of the IEEE Symp. on Information Visualization (InfoVis 2004). 17-24.

[4] Blackwell, A.F., Engelhardt, Y. 1999. A meta-taxonomy for diagram research. Anderson, M., Meyer, B., Olivier, P. (Eds.), *Diagrammatic Representation and Reasoning*. London: Springer-Verlag. 109-123.

[5] Duke, D.J., Brodlie, K.W., Duce, D.A. 2004. Building an ontology of visualization. *Proc. of the Conf. on Visualization '04*. 589.7

[6] Wehrend, S., Lewis, C. 1990. A problem-oriented classification of visualization techniques. *Proc. Of the 1st Conference on Visualization '90*. 139-143.

[7] Yu, C.H., Behrens, J.T. 1995. The alignment framework for data visualization: relationships among research goals, data types, and multivariate visualization techniques. Paper presented at the *Annual Meeting of the American Psychological Society*, San Francisco, CA.

[8] Shneiderman. B. 1996. The eyes have it: a task by data type taxonomy for information visualizations. *Proc. Of the 1996 IEEE Symp. On Visual Languages.* 336.

[9] Zhou, M.X., Feiner, S.K. 1998. Visual task characterization for automated visual discourse synthesis. *Proc. of the CHI '98*. 392-399.

[10] Wenzel, S., Bernhard, J., Jessen, U. 2003. A taxonomy of visualization techniques for simulation in production and logistics. Chick, S., Sánchez, P.J., Ferrin, D., Morrice, D.J. (Eds.), *Proc. of the 2003 Winter Simulation Conference*. 729-736.

[11] Tory, M. Möller. T. 2004. Rethinking visualization: a high-level taxonomy. *Proc. of the IEEE Symp. on Information Visualization 2004 (InfoVis 2004)*. 151-158.

[12] Chi, E.H. 2000. A taxonomy of visualization techniques using the data state reference model. *Proc. of the IEEE Symp. on Information Visualization 2000 (InfoVis 2000)*. 69-76.

[13] Bertin, J. 1983. *Semiology of Graphics*. Madison, Wisconsin: University of Wisconsin Press.

[14] Mackinlay. 1986. Automating the design of graphical presentations of relational information. *ACM Trans. On Graphics*. **5**(2): 110-141.

[15] Rogowitz, B.E., Treinish, L.A. 1993. An architecture for rule-based visualization. *Proc. Of the 4th Conf. on Visualization '93*. 236-243.

[16] Zhang, J. 1996. A representational analysis of relational information displays. *Int. J. Human-Computer Studies*. **45**: 59-74.

[17] Zhou, M.X., Ma, S., Feng, Y. 2002. Applying machine learning to automated information graphics generation. *IBM Systems Journal.* **41**(3): 504-523.

[18] Zhou, M.X., Chen, M., Feng, Y. 2002. Building a visual database for example-based graphics generation. *Proc. of the IEEE Symp. on Information Visualization 2002 (InfoVis 2002).* 23-30.

[19] Lohse, J., Rueter, H., Biolsi, K., Walker, N. 1990. Classifying visual knowledge representations: a foundation for visualization research. *Proc. of the 1st Conf. on Visualization '90*. 131-138.

[20] Velez, M.C., Silver, D., Tremaine, M. 2005. Understanding visualization through spatial ability differences. *Proc. of IEEE Visualization (Vis 2005).* 511-518.

[21] Santella, A., DeCarlo, D. 2004. Robust clustering of eye movement recordings for quantification of visual interest. In: *Eye tracking research and applications (ETRA) Symp. 2004.* 27-34.

[22] Itti, L., Baldi, P. 2006. Bayesian surprise attracts human attention. In: *Advances in Neural Information Processing Systems (NIPS*2005).* **19**: 1-8. Cambridge, MA: MIT Press.

[23] von Ahn, L., Dabbish, L. 2004. *Proceedings of the SIGCHI conference on human factors in computing systems (CHI 2004).* 319-326.