# Gaze informed View Management in Mobile Augmented Reality

**Ann M. McNamara**
Department of Visualization
Texas A&M University
College Station,
TX 77843 USA
ann@viz.tamu.edu


CHI 2013 Workshop on "Gaze Interaction in the Post-WIMP World", April 27, 2013, Paris, France.

## Abstract
Augmented Reality (AR) systems provide an enhanced vision of the physical world by integrating virtual elements, such as text and graphics, with real-world environments. AR allows us annotate the physical world with virtual information to enhance our understanding, enjoyment and usefulness of our immediate surroundings. A fundamental problem is the optimal integration of real and virtual elements, and provide a seamless user-integration paradigm. This problem is amplified with mobile AR platforms due to the implicit reduction in available screen real estate. This paper describes early research intended to develop principled algorithms for optimized integration of real and virtual elements in mobile AR based on user-attention.

## Author Keywords
Guides, instructions, author's kit, conference publications Mobile Augmented Reality, Gaze Interation, Eye Tracking, Visual Clutter

## ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous. H.5.2 User Interfaces: Evaluation/methodology, H.1.2 User/Machine Systems: Human information processing

## Introduction

Augmented Reality (AR) systems provide an enhanced vision of the physical world by integrating virtual elements, such as text and graphics, with real-world environments. AR allows us annotate the physical world with virtual information to enhance our understanding, enjoyment and usefulness of our immediate surroundings. The advent of affordable mobile technology has sparked a resurgence of interest in mobile AR applications. A fundamental problem is the optimal integration of real and virtual elements. Integrating real and visual elements poses a challenge because augmented elements, or labels, are contextually linked to real world objects or locations. To ensure the correct associations between virtual elements and real objects the augmented elements must be placed in the vicinity of the object it describes. Enforcing these spatial associations can lead to undesirable results: labels can overlap each other rendering them unreadable and, labels can obscure real world objects that are relevant to the user. Optimal placement of labels is an active area of research. This research will evaluate the benefits of identifying where the user is looking and placing information in that location only. Gaze interaction will be used to determine both where the user is looking and to guide optimal placement by only displaying virtual elements associated with the real objects that the user attends to. This research develops principled algorithms for optimized integration of real and virtual elements in mobile AR based on user-attention.

## Visual Clutter in AR

Augmented elements such as text annotations can contribute to *visual clutter* [8, 5, 6, 7, 23, 12]. In environments with many elements there is the distinct possibility that elements will *overlap* with each other, or *obscure* important information in the scene [6]. There has been considerable research done on the optimal placement of annotations in a scene, which turns out to be an NP-complete problem (the number of alternative placement areas grows exponentially with the number of objects in the scene) [7, 18, 10, 19]. While several solutions have been proposed including greedy algorithms, cluster-based methods and screen subdivision methods [1, 17, 21, 22], no existing research that uses gaze information about where the user is looking in the scene to place augmented elements in the scene. This will it will make element placement more effective, and ensure that placement does not obscure important information. A variety of approaches exist to reduce the number of labels by filtering information based on properties of the task and user proximity [11, 12, 13, 14] . Bell et. al describe an alternative strategy that places labels in "non-interesting" regions and keeps track of the free areas of the display [3, 2]. Some researchers have used the properties of labels to determine a placement strategy, for example color, transparency and readability [4, 24]. Henderson et al. found that clutter correlates with search performance in a real-world scenes. Furthermore, they provided evidence that clutter also predicts eye movement characteristics during real-world search [9, 20]. These data converge with those presented by Rosenholtz et al. [17] in showing that an image-based proxy for search set size can be related to search performance in real-world scenes.

## A New Approach: Gaze Informed View Management

The work will develop a subdivision scheme to divide the real scene into smaller regions or *cells* and associate the labels for the anchors in that region of each cell. This scheme will depend on the visual richness of the real scene. The work will proceed by first dividing the screen into individual cells, these may be quadrants, octants or even higher resolution. The resolution will depend on the visual clutter in the real scene; in homogenous scenes where there is little variability, there may be no need to subdivide. In visually rich scenes a higher number of smaller cells should work better. Existing measures of visual clutter (for example [16, 17]) to select the appropriate number of cells. Labels will then be sorted into these cells. Labels which annotate anchors that occupy multiple cells will need to be duplicated into the cells occupied by that anchor. All of these algorithms will be optimized for fast processing on the mobile device. During navigation the gaze position on the mobile device will be communicated to the AR application either via the internal hardware on the device or via bluetooth if the eye-tracking device is a separate component. The algorithm will check which cell the gaze position falls in and then display the labels for that cell. As the user re-directs her gaze and gaze enters a new cell the current labels will gracefully degrade while the new labels are presented. It is important they fade gradually rather than simply disappear to ensure smooth transition of attention from one anchor to the next in order to minimize any distracting phenomena introduced by toggling label visibility.

## Mechanisms

Visual clutter actually suppresses the brain's responsiveness. Kastner et al. hypothesized that by focusing its attention on just one stimulus, the brain cancels out the suppressive influence of nearby stimuli. In this way, it enhances information processing of the desired stimulus. It is clear to see the advantage of such a strategy for mobile AR. Here we use eye-tracking to determine the relevant stimulus and display it, while hiding others. A **label** is semantic information that can be attached to a **Point of Interest** (POI) or **anchor**. A label is typically displayed in an axis-aligned bounding box which is known as the **label box**. A problem in mobile AR is how “best” to place these label boxes. *Best* can be interpreted in different ways but in general is taken to mean that the boxes a) do not overlap, b) clearly annotate their anchor and c) obscure as little of the real world as possible. Many strategies have been researched to determine the optimal placement, size and timing of label boxes [15, 12, 11, 13, 14]. First we consider using gaze position as the determining factor. There are clear benefits to doing so. By presenting label boxes only in the region the user looks minimizes the number of elements to display thereby minimizing the risk of overlap. Without risk of overlap clear associations with anchors can be ensured and once the user re-directs their gaze the label boxes in the unattended region do not need to be displayed, thereby minimizing the proportion of the real world that is obscured. **Figure 1** (center) shows a likeness of how this would look. Clearly the view in the center minimizes overlap, provides clear annotation and obscures less of the real world. The main uncertainty of this approach lies in the level of distraction that may be introduced by eliminating and (re)introducing labels dynamically as the viewer gazes over an image. The proposed solution will use a combination of screen position and eye-movement to ensure that label placement does not become distracting.


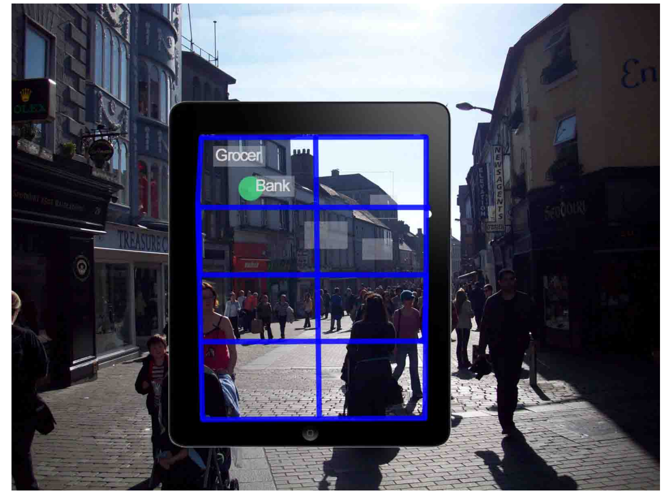


**Figure 1:** Many current Mobile AR applications present all available information based on location & camera direction. This can lead to visual clutter as shown in the image on the top, even though in this example there is a rather conservative number, seven, of labels. It would be neater to present information for the buildings the user is looking at. A "mock-up" of this is shown in the center, there is no distracting information, simply data for where the user is looking which makes for a much cleaner view. A sketch over (right) shows the potential placement of "cells" and eye-position.

## Conclusions

This paper describes initial research on principled algorithms for the optimized integration of real and virtual elements in mobile AR based on user-attention, specifically user gaze. Placement is based on user attention. This will minimize visual clutter and lead to improve view management systems for mobile AR platforms.

## References


[1] Azuma, R., and Furmanski, C. Evaluating label placement for augmented reality view management. In *ISMAR*, IEEE Computer Society (2003), 66–75.

[2] Bell, B., Feiner, S., and Hollerer, T. View management for virtual and augmented reality. ACM Press (2001), 101–110.

[3] Bell, B., Feiner, S., and Höllerer, T. Information at a glance. *IEEE Comput. Graph. Appl. 22* (July 2002), 6–9.

[4] Bernard, M. L., Chaparro, B. S., Mills, M. M., and Halcomb, C. G. Comparing the effects of text size and format on the readibility of computer-displayed times new roman and arial text. *Int. J. Hum.-Comput. Stud. 59* (December 2003), 823–835.

[5] Bravo, M. J., and Farid, H. Search for a category target in clutter. *Perception 33*, 6 (2004), 643–652.

[6] Bravo, M. J., and Farid, H. A measure of relative set size for search in clutter. *Journal of Vision 7*, 9 (2007), 1054.

[7] Bravo, M. J., and Farid, H. A scale invariant measure of clutter. *Journal of Vision 8*, 1 (2008).

[8] Ellis, G., and Dix, A. A taxonomy of clutter reduction for information visualisation. *IEEE Transactions on Visualization and Computer Graphics* (2007).

[9] Henderson, J. M., Chanceaux, M., and Smith, T. J. The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements. *Journal of Vision 9*, 1 (2009).

[10] Menozzi, M., and Koga, K. Visual information processing in augmented reality: Some aspects of background motion. *Swiss Journal of Psychology 63*, 3 (2004), 183 – 190.

[11] Peterson, S. D., Axholt, M., Cooper, M., and Ellis, S. R. Evaluation of alternative label placement techniques in dynamic virtual environments. In *Proceedings of the 10th International Symposium on Smart Graphics*, SG '09, Springer-Verlag (Berlin, Heidelberg, 2009), 43–55.

[12] Peterson, S. D., Axholt, M., Cooper, M., and Ellis, S. R. Visual clutter management in augmented reality: Effects of three label separation methods on spatial judgments. *3D User Interfaces 0* (2009), 111–118.

[13] Peterson, S. D., Axholt, M., and Ellis, S. R. Comparing disparity based label segregation in augmented and virtual reality. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, VRST '08, ACM (New York, NY, USA, 2008), 285–286.

[14] Peterson, S. D., Axholt, M., and Ellis, S. R. Label segregation by remapping stereoscopic depth in far-field augmented reality. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '08, IEEE Computer Society (Washington, DC, USA, 2008), 143–152.

[15] Polys, N. F., Kim, S., and Bowman, D. A. Effects of information layout, screen size, and field of view on user performance in information-rich virtual environments. In *Proceedings of the ACM symposium on Virtual reality software and technology*, VRST '05, ACM (New York, NY, USA, 2005), 46–55.

[16] Rosenholtz, R., Li, Y., Jin, Z., and Mansfield, J. Feature congestion: A measure of visual clutter. *Journal of Vision 6*, 6 (2006), 827.

[17] Rosenholtz, R., Li, Y., and Nakano, L. Measuring visual clutter. *Journal of Vision 7*, 2 (2007).

[18] van den Berg, R., Cornelissen, F. W., and Roerdink, J. B. T. M. A crowding model of visual clutter. *Journal of Vision 9*, 4 (2009).

[19] Verghese, P., and McKee, S. P. Visual search in clutter. *Vision Research 44*, 12 (2004), 1217 – 1225. Visual Attention.

[20] Wichmann, F. A., Kienzle, W., Schlkopf, B., and Franz, M. Non-linear system identification: Visual saliency inferred from eye-movement data. *Journal of Vision 9*, 8 (2009), 32.

[21] Wither, J., DiVerdi, S., and Hollerer, T. Annotation in outdoor augmented reality. *Computers and Graphics 33*, 6 (2009), 679 – 689.

[22] Wither, J., Tsai, Y.-T., and Azuma, R. Indirect augmented reality. *Computers and Graphics 35*, 4 (2011), 810–822.

[23] Wolfe, J., and Horowitz, T. Visual search. *Scholarpedia 3* (2008).

[24] Zhuang, X., and Papathomas, T. V. Cue relevance effects in conjunctive visual search: Cueing for location, color, and orientation. *Journal of Vision 11*, 7 (2011).