Winter 12-5-2005

Research on Issues Related to Virtual Reality Representation in Online Shopping System –Experimental Study on Spatial Location Relations among Objects Based on Visual Attention Theory

Jiangyu Li

Haibo Yu

Research on Issues Related to Virtual Reality Representation in Online Shopping System –Experimental Study on Spatial Location Relations among Objects Based on Visual Attention Theory

Li Jiangyu *, Yu Haibo **
* School of Economic & Business Administration,
Beijing Normal University, China, 100875
Telephone: 86-010-62261810, Fax numbers: 86-010-58801867
e-Mail: li_jy@126.com
** The Institute of Psychology
Chinese Academic of Sciences, China, 100101

Abstract: This article has studied influences of three factors namely flashing (FF), consistency (CF) and distance (DF) to spatial location array of commodity objects in online shopping system. The findings show that reaction time (RT) of subjects in still representation is shorter than that in live representation; faster in the setting of being inconsistency between commodities and their text description than that of being consistency; in the condition of both Flashing (FSH) level and Consistency (CON) level, Un-flashing (UNFSH) level and Inconsistency (INCON) level, subjects RT in the setting of commodity object being farther from its description is shorter than that in being near. The research finds no discrepancies that three factors have any impact on subjects' accuracy rate. Further analysis finds that physical distance of commodities plays a major role in affecting spatial location array of objects, while conceptual distance ranks on the second place. Location-based visual attention has the biggest impact on spatial location relation of virtual reality setting, and object-based visual attention plays a second largest impact. The impact would be highly impressive when either physical distant or conceptual distance is conformed to experiences in real life, Singleton detection mode will play a role at the situation when coincidence of the said phenomenon and real life experience is less, which means distinct flashing (FSH) will lead to better effect at this particular situation, otherwise it is worse.

Keywords: E-commerce Management, Internet Marketing & Advertising, Attention theory, virtual reality, location relations, shopping system.

I. Introduction

Some research finds that virtual reality representations of commodity information in online shopping systems will encourage buyers to reach an positive assessment on commodity information, and further inspires their desire of buying behavior[1]. This finding and further studies on affecting factors of VR representation to human information behavior or buying behavior in e-commerce websites grab a lot of attention from scholars and researchers, and ongoing studies are implemented, for example: user structure pattern of virtual experience in VR environment; the relation between structure pattern and users' direct or/and indirect experience[2]. Another finding related to the means of media representations shows that VR representations have more advantages in improving performance efficiency specially when buyers are familiar to shopping systems or face higher frequencies of buying activities than Text representation, therefore, object information represented in VR plays a significant role on the success of online shopping systems[3].

In terms of representation on specific blocks of the interface and their relations, conventional researches focus more on micro-cognition elements such as interface layout, the characteristics of font and color and their relations with viewing behavior [4], which also are the issues of web interface our attention should be paid to, not limited in this, a further intensive attention should be paid to the different features and models of these elements in web environment [5], while some results might not be well applied to object representation in VR, and existing researches on this issue are much more scarce so far.

VR representation, in a way, can be interpreted as a means of presenting spatial relation which is largely based on intuition. It can easily lead to misunderstanding due to its capacity of sending diverse information simultaneously, so more aid such as text description are employed to reduce misunderstanding and non-clarity while user interacting with spatial activities in VR, so user can differentiate specific objects from its background or other elements unrelated to establish information cognition and activities pattern in VR environment[6]. Some researchers, from another point of view, have studied shopping design in real life and its applicability at VR stores, the impact of VR shop individuality on user's behavior[7]. In VR shopping environment, the issues of information behavior and location relations among objects still need more attention. This paper mainly studies the impact of attention theory on objects and object relations in VR setting.

II. Visual Attention Theory

Attention, in terms of the cause, can be divided into two classes: passive attention and active attention, some researchers also name them as stimuli-driven attention or

Proceedings of the Fifth International Conference on Electronic Business, Hong Kong, December 5-9, 2005, pp. 436 - 440.

target-driven attention [8], yet two classes of attention are not mutually exclusive, sometimes, the development or controlling of the attention is worked jointly by the two. Passive attention often involves two modes: singleton detection mode and feature search mode. Singleton detection mode refers to the distinct difference of the stimuli (stimuli refers to the means of commodity representation in online shopping system) from the background. Feature search mode refers to the consistency of stimuli features and task features. Therefore, visual attention can be led to specific locations because of all said features.

Attention selection to the stimuli consists mainly of two categories: location-based visual attention and object-based visual attention. Location-based attention believes that attention is influenced by spatial array-format of objects. The selection accuracy and speed of subjects to searching objects are better and faster when objects appear near than they appear far. Other research findings support that the selection of searching objects are influenced by not only spatial location, but also location relations among objects. That is what we called object-based attention. Both two categories of attention will be explicated further by demonstrating an example. To explain location-based visual attention, we present a figure (1) of rectangle. An object can appear at each corner of the rectangle randomly. Subjects are asked to report another object's identity which is associated with the object as soon as it appears. The speed and accuracy of subject selection is faster and better when two objects are spatially set closer than they are far[9]. To explain object-based visual attention, we present a figure (2) of two same sized rectangles, the appearance of an object (named A here) appears with eighty percentage at fixation, and Bs either on left or on right side of A appear with ten percentage at any possible places, but the distance to A from either B remain the same. Subjects are asked to report as soon as either B appears, and subjects RT when B and A are in the same rectangle is shorter than that when B and A are in two separate rectangles[10,11]. It also reveals that simple detection task is more applicable to structure locations of objects, while objects identity task will require lots of attention paid to objects representations[12].

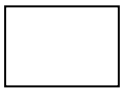

Figure 1

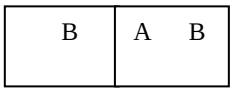

Figure 2

The research on visual searching finds that the speed of attention relocated to other objects and duration the attention kept on the objects have strong correlation with the relations among objects. The bigger discrepancies among objects exist, the longer the attention would be paid[13].

In VR representation of online shopping systems, users need experience different shopping sites and differentiate specific shopping objects while touring information space. User attention will be constantly in the mode of stimuli or targets or the rotation of the two resulted from the changes of sites and man's cognition. Users will go through two subsequent stages: detection task and objects identity task. Detection task involves more direct experiences and attention to the location and space of the setting, while objects identity task involves more attention to specific object, its shape, its characteristic etc. To some extent, specific objects in VR shops should be distinguished from the background or other objects related since they have direct impact on buying decision and information activities. The demonstration of VR space can be the specific location relations of objects, the specific features of objects, or the combination of the two. The representation by either way will become clear in user mind when they catch more user attention successfully. In most case, the information related to objects should be placed in one sphere instead of two separate ones.

III. Hypothesis

Based on visual attention theory explicated previously, this article mainly focuses on spatial location relations among objects in VR setting, studies on effects of three factors to spatial location relations among objects have been made in VR online shopping system and three hypothesis are going to be validated. Namely, whether the result of task performed will be better? if 1. two objects sharing same characteristics are placed at a consistent location array. 2 two objects linking each other logically are near or placed in the same spatial range. 3 targeting object is distinct from its objects' background. In this article, objects and their description can be interpreted as two objects sharing same characteristics. Distance being either far or near can be interpreted as two objects being close or far. Objects are flashing will be interpreted as they are distinct from their background or vice versa.

IV. Method

Subjects：80 sophomores from Beijing Forestry University (52 males, 28 females) are with little experiences of using computer, their average age is 20.13.

IV. 1 Experimental Materials：

Eight VR settings has been employed, and each setting consists of three parts of shelf shaped like "⌐¬" , each shelf has four booths divided by lines and placed at four corners of each shelf. Each commodity at booth accompanies a description indicating the nature of goods. Eight settings are different because of the array of three factors.

Flashing (FSH) refers to the change of commodity objects becoming either larger or smaller. Un-flashing (UNFSH) refers to the size of commodity objects remaining the same. Consistency (CON) refers that location relations of commodity objects and its description remain the same at

booth, namely it means commodity object always appears at lower left corner of the booth, while its description appears at upper right corner of the booth. Inconsistency (INCON) refers that location relations of commodity object and its description appear randomly at different places at booth. The definition of farther (FNR) or nearer (NR) depends on whether commodity object and its description are blocked off by a line while physical distance remains the same. If it is a line between commodity object and its description, it means FNR; if not, it means NR. Subjects can only view one booth containing four commodities while their entering into shopping settings. Subjects must "go around" if they want to view more booths or commodities. "Go around" and final selection of commodity objects here must be facilitated by moving mouse.

IV. 2 Experimental Task

A target object appears randomly out of 48 objects in the shopping center. Subjects navigate through the specific setting by means of the mouse. Subjects then click the "SELECT" button upon arrival at the target object, indicating completion of one task. The setting will disappear after task completion, and next target object will appear, which indicates the start of second task to be executed. Apart from objects, all commodities and their description in the setting appeared randomly.

IV. 3 Experimental Design

Three-factor within-subjects design (2x2x2) generated eight treatments which correspond with eight virtual reality settings. Eighty subjects were randomly grouped into eight teams evenly and assigned a treatment.

Independent Variables: Three factors and two levels in each factor, namely flashing factor (FF)-(flash vs un-flash), consistency factor (CF)-(con-incon) and distance factor (DF)-(farther vs nearer)

Dependent Variables:

1. Reaction Time (RT): The time a subject spends completing a task. Time begins when an object appears, and ends when the subject locates the object and clicks "SELECT."
2. Accuracy: The percentage of all correct object-searching tasks performed by each subject. It is resulted from the ratio between the number of correct tasks done by subjects and the number of whole tasks to be done.

Controlled Variables:

Time Delay (TD): refers to about nine seconds lapsed while each scene appears on the screen of the computer.

IV. 4 Experimental Procedures

In each treatment (eight settings) subjects were first asked to read task instructions, and then perform a practice task in an attempt to let subjects be familiar to and understand the whole task to be implemented (practice task content was unrelated to experimental task content). Subjects practiced ten search tasks before proceeding to the experimental tasks. As soon as a random target object appeared, subjects clicked "START", then subjects navigated through shopping setting until the correct object was located, completing the search activity by clicking "SELECT." A new random target object then appeared, initiating the second task. The search process was repeated twenty times by each subject, with the computer monitoring the whole process and recording all activities.

V. Experimental Results

V. 1 Reaction time (RT) under different treatments

Table 1: Reaction time under different treatments

FF	CF	DF	Mean(ms)	Deviance	Subjects
FSH	CON	FNR	13811	4728	8
		NR	13910	3135	9
		Total	13864	3834	17
	INCON	FNR	11975	1385	10
		NR	10976	2356	11
		Total	11452	1975	21
	Total	FNR	12791	3332	18
		NR	12296	3050	20
		Total	12531	3153	38
UNFSH	CON	FNR	16103	6936	11
		NR	10848	3535	11
		Total	13476	6008	22
	INCON	FNR	8922	1618	10
		NR	9551	2373	11
		Total	9255	2024	21
	Total	FNR	12684	6224	21
		NR	10200	3012	22
		Total	11413	4956	43
Total	CON	FNR	15138	6064	19
		NR	12226	3628	20
		Total	13645	5116	39
	INCON	FNR	10449	2145	20
		NR	10263	2420	22
		Total	10352	2267	42
	Total	FNR	12733	5036	39
		NR	11197.9860	3175.2211	42
		Total	11937.2331	4219.9678	81

Flashing factor=FF
Consistency factor-CF
Distance factor=DF

In general, RT used by subjects in FSH（X=12530.7288 ms）is less than that in UNFSH（X=11412.7485ms）; RT in CON（X=13644.7608ms）is less than that in INCON（X=10351.6716ms）; RT in FNR（X=12733.3453ms）is less than that in NR（X=11197.9860ms）. These RT scores are then log-transformed to normalize the data（RT—>Log（RT+1）, then conduct ANOVA analysis (GLM General Factorial, SPSS for Windows 8.0).

The results appear graphically in Figure 3 and Figure 4. overall differences among eight treatments are highly significant（$F(7,73)=3.895, p<0.01$）. The main effect for FF（$F(1,79)=5.478, p<0.05$）and for CF（$F(1,79)=12.827,$

p<0.01）both are significant. The three-way interaction among three factors is significant （ F(1,79)=4.162, p<0.05）, however, there is no main effect for DF (F(1,79)= 1.726,p=0.193）. There is also no significant two-way interaction between FF and CF （ F(1,79)=1.085, p=0.301）, between FF and DF （F(1,79)=0.643 p=0. 425）or between CF and DF（F(1,79)=0.951, p=0.333）.

Figure 3 and Figure 4 shows that subjects RT in CON level is longer than that in INCON level at each level of both DF and CF. In other conditions except FNR level and CON level under flashing factor, subjects RT in FSH level is longer than that in UNFSH level. In CON level, subjects RT in NR and UNFSH condition is shorter than that in other conditions. In INCON level, the speed in FNR and UNFSH condition is faster than that in NR and UNFSH while the speed in NR and FSH condition is faster than that in FNR and FSH.

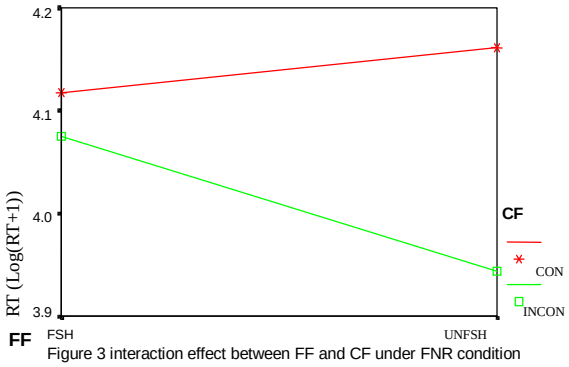

Figure 3 interaction effect between FF and CF under FNR condition

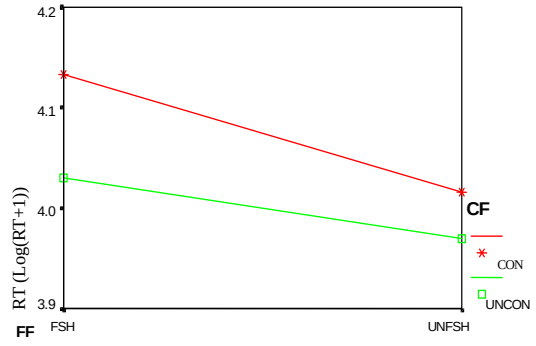

Figure 4 interaction effect between FF and CF under NR condition

V. 2 The accuracy under different treatments

Statistics analysis（GLM General Factorial in SPSS for Windows 8.0 ）shows that overall differences among treatments are not significant （F(7,73)=1.466,p=0. 931）. The main effect for FF（F(1,79)=3.029,p=0. 086）, for CF （ F(1,79)=1.194,p=0.278 ）and for DF （ F(1,79)=2.644, p=0.108） all are not significant. There is also no significant two-way interaction between FF and CF（F(1,79)=0.042, p=0.837 ） , between FF and DF （ F(1,79)=2.699, p=0.105 ） or between CF and DF （ F(1,79)=2.699, p=0.105）. The three-way interaction among three factors is also not significant （F(1,79)=0.196, p=0.660）.

VI. Analysis and Discussion

In all, researching findings and hypnosis can't reach an agreement, and the impact of DF still remains unclear. The interaction of three-way factors will come to different results under different conditions.

The definition on CON level of consistency factor is that commodity object and its description are placed at diagonal corner in rectangle, actually, which makes the two being farthest physically. The definition of distance factor mainly refers to the conceptual distance between commodity objects and its description, and physical distance of two levels (FNR and NR) in distance factor is the same. It will be easier to understand that consistency factor involves physical location relation of the attention, while distance factor involves conceptual distance of the attention, and conceptual distance is a more objects-oriented process paid by attention, which is conformity with the said analysis. INCON level makes commodity object and its description nearest physically, which is coincidently agreement with the setting at real life stores where tags are placed closely to commodities, therefore the similarity in location relations plays the greatest impact on subjects cognition, and help subjects identify objects at shortest time. CON level tells another opposite story at any conditions, that is results in Con level at any conditions are worse than that in INCON level.

Based on attention theory, the speed of human apperceiving live objects is faster than that of still objects if moving objects appear distinctly from the background. In CON and FNR condition, response speed of subjects encountering flashing objects is faster than that of un-flashing objects, while in other conditions, opposite results are found, which is probably because subject mentality is required to identify objects description clearly first and give a judge whether they choose right targets, while flashing objects more easily grab subject's instinct attention. The process of this specific task would follow the pattern: subjects' attention moving to flashing objects first—then to objects description to identify the correction—then to search target object after judgment, the whole process slows down the cognition speed. If objects are not flashing, the attention of subjects will be paid directly to objects, which accelerates cognition speed. Therefore results of farther level in FSH and CON condition is better than that of nearer level because being farther further reduced the impact of distance. Apart from this, in other conditions, other factors affecting cognition is greater than flashing. In CON and FNR level, it can be interpreted as the worst cognition mode since both

physical and conceptual distance are fair faraway. Cognition results are better in other conditions because physical distance and conceptual distance are near. In other conditions, subjects RT in flashing level is shorter than that in un-flashing.

Based on the studies and experiments, the conclusion drawn is that physical-distance influence on location relations among objects is biggest, conceptual distance is on second place. In terms of attention theory, the influence on location relations by location –based attention is bigger than that by object-based attention. Performance results are better when the said phenomenon is coherent with experiences in real life. Singleton detection mode will play a role at the situation when coincidence of the said phenomenon and real life experience is less, which means distinct flashing will lead to better effect at this particular situation, otherwise it is worse. It helps us better understand and establish VR shopping settings, to use more location relations in physical terms to represent the relations of objects instead of conceptual terms; to place commodity objects which are similar to each other at more close locations as so to reveal more information about commodities, however location relations of three factors determining buying decision need be further studied.

References

[1] Alice Richmond. Enticing online shoppers to buy – A human behavior study. *Proceedings of the fifth international World Wide Web conference on Computer networks and ISDN systems*. Amsterdam: Elsevier Science Publishers, 1996, 1469-1480

[2] Hairong Li, Terry Daugherty, Frank Biocca. Characteristics of Virtual Experience in Electronic Commerce: A Protocol Analysis. http://elab.vanderbilt.edu/research/papers/pdf/ manuscripts/ im.pdf, 2005-05-20.

[3] Li Jiangyu. Shopping-online: the Effect of Media Representation Patterns on Human Information Behavior. In: Li Qi. *Selected Proceedings, the Second China and U.S. Advanced Workshop in Electronic Commerce 2004*. Chendu, Southwest University of Finance & Economic Press, 2004, 331-337

[4] Thomas S. Tullis. User Interface Design, *Handbook of Human-Computer Interaction*. Amsterdam, North-Holland, 1988, 377-411

[5] Lynch P. J., Yale. University C/AIM WWW Style Guide, http://info.med.yale.edu/caim/ StyleManual_ Top.HTML, 1995

[6] Li Yi-xiong & Chen Ze-min，*Cognition Psychology*，Psychology Press Ltd. Co. Taiwan, 1990,177-188

[7] Luca Chittaro, Roberto Ranon. New Directions for the Design of Virtual Reality Interfaces to E-Commerce Sites[EB/OL], http://hcilab.uniud.it/ publications/2002-04/5.pdf, 2005-05-20.

[8] Howard E. Egeth & Steven Yantis. “VISUAL ATTENTION: Control, Representation, and Time Course，” *Annual Review of Psychology*. 1997. 48, 269-297.

[9] Hoffman, JE & Nelson, B. “Spatial selectivity in visual search,” *Perception and Psychophysics*, l98l, 30, 283-290.

[10] Behrmann, M, Zemel, RS, & Mozer, MC. “Object-based attention and occlusion: Evidence from normal subjects and a computational model,” *Journal of Experimental Psychology: Human Perception and Performance*, 1998, 24(4), 1011-1036.

[11] Egly R, Driver J, Rafal RD.. Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. *J. Exp. Psychol.: Gen.* 1994,123,161–177

[12] Vecera, SP, & Farah, MJ. “Does visual attention select objects or locations ?”*Journal of Experimental Psychology: General*, 1994, 123(2)，146-160.

[13] Moore, C. M., Egeth, H., Berglan, L. R., & Luck, S. J. “Are attentional dwell items inconsistent with serial visual search ?” *Psychonomic Bulletin & Review*, 1996, 3, 360–365.