Multimed Tools Appl
DOI 10.1007/s11042-015-2804-1

Method and experiments of subliminal cueing for real-world images

Tai-Hsiang Huang[1] **· Su-Ling Yeh**[2] **· Yung-Hao Yang**[3] **·**
Hsin-I Liao[3] **· Ya-Yeh Tsai**[3] **· Pai-Ju Chang**[3] **·**
Homer H. Chen[4]

Received: 3 February 2015 / Revised: 28 April 2015 / Accepted: 1 July 2015
© Springer Science+Business Media New York 2015

Abstract Unconscious attention shift triggered by a subliminal cue has been shown to be automatic; however, whether it can be brought into effect for images of real-world scenes remains to be investigated. We present a subliminal cueing method that flashes briefly a visual cue before presenting a real-world image to the viewer. The effectiveness of the method is verified by experiments using three types of cues (spatial cue, face cue, and object cue) of varied durations. Results show that depending on the cue type, the viewer's visual attention is

✉ Tai-Hsiang Huang
tshuang1983@gmail.com

Su-Ling Yeh
suling@ntu.edu.tw

Yung-Hao Yang
yunghaoyang@gmail.com

Hsin-I Liao
irisliao0111@gmail.com

Ya-Yeh Tsai
r00227103@ntu.edu.tw

Pai-Ju Chang
b97207003@ntu.edu.tw

Homer H. Chen
homer@ntu.edu.tw

1 Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617 Taiwan, Republic of China

2 Department of Psychology, Graduate Institute of Brain and Mind Sciences, and Neurobiology and Cognitive Science Center, National Taiwan University, Taipei 10617 Taiwan, Republic of China

3 Department of Psychology, National Taiwan University, Taipei 10617 Taiwan, Republic of China

4 Department of Electrical Engineering, Graduate Institute of Communication Engineering, and Graduate Institute of Networking and Multimedia, National Taiwan University, Taipei 10617 Taiwan, Republic of China

Published online: 05 August 2015

directed to the cued visual hemifield or the cued location without engaging the viewer's awareness. The experiments demonstrate that a brief subliminal cue presented prior to the color image of a real-world complex scene can attract human visual attention. The method is useful for many applications that require efficient, unresisting attention shift to a target image area.

Keywords Visual attention · Eye movements · Eye tracking experiment · Subliminal cue · Human visual system

1 Introduction

Visual attention is an important characteristic of the human visual system (HVS). It helps our brain to filter out excessive visual information and enables us to focus on a particular object or region of interest [13, 14]. Many biologically inspired algorithms attempt to mimic this fascinating characteristic of HVS for applications in image processing [30, 57], video coding [50], biomedical imaging [24], and video quality assessment [11], to name a few.

While considerable efforts have been directed towards predicting the focus of human visual attention in the image processing community [21, 29], little about how to actively direct human visual attention to a target area has been investigated. In practice, it is extremely easy to arouse viewer's attention by deliberately changing, for example, the color of the target region, but in many applications such explicit guidance is prohibited because it would ruin the aesthetic appeal of the image and spoil the viewing experience. Moreover, an explicit change in the visual display may instead drive the human visual attention away from the planned region [27].

To avoid the problems mentioned above, implicit or nearly implicit attention guidance is preferred. In this work, we are concerned with the guidance of human visual attention without involving the viewer's awareness. Such a technique can make, for example, a digital signage advertisement more effective because it pulls and holds the viewers' attention while avoiding top-down controlled resistance in human perception as discovered in scientific research [31, 37, 53]. Media design and other applications such as digital art and education may also benefit from the technique [44]. For example, it has been shown that implicit cues can enhance the learner's inductive reasoning abilities and increase learner's performance and intuition in logic-based problem-solving tasks [6]. It has also been shown that athletes perform better when exposed to subliminal cues [3] and that human computer interaction can be more pleasing and effective when subliminal cues are involved [45]. It is our interest to investigate an effective presentation of implicit cues for real-world applications.

Implicit information is known to massively affect human perception and behavior [32]. Although a weak visual input may not lead to conscious perception, it affects the visual response. It has been shown that subliminal stimuli, though not consciously perceived, can affect human's conscious perception and behavior [8, 9, 18, 38]. For example, Karremans et al. found that subliminal presentation of a brand name of drink can affect thirsty participants' choice of that brand [31], Yang et al. and Yeh et al. found that semantic information can be extracted from invisible or unrecognizable words [62, 64], and Chou et al. showed that an invisible object can guide viewer's visual attention just as a visible object does [9]. The burgeoning evidence of the powerful influence of subliminal stimuli in psychology studies motivates us to develop a subliminal cueing method to bring implicit human visual attention guidance to life.

Based on our previous work [20], which shows that a flashed visual cue can subliminally direct human visual attention, we further explore implicit visual attention guidance for real-world applications and investigate the temporal and spatial efficiency of local, face, and object cues for implicit visual attention guidance in this work.

The remainder of the paper is organized as follows. Section 2 provides a brief literature review on visual attention guidance. Section 3 describes a system for generating attention-guiding images and the design considerations of the system. Section 4 describes the eye tracking apparatus used in our experiments to collect visual attention data and the details of four performance evaluation experiments. A further discussion of the experimental results is given in Section 5. Finally, the concluding remarks are drawn in Section 6.

2 Review

It is well recognized in psychology that visual attention can be oriented in two different ways: voluntary and involuntary. Voluntary attention refers to orienting towards stimuli relevant to the behavioral goal of a viewer. For example, a driver would attend to pedestrians, cars, traffic lights, etc., while driving since these stimuli are relevant to the behavioral goal (i.e., safe driving) of the driver. Involuntary attention refers to orienting towards stimuli that are irrelevant to the behavioral goal of a viewer but are irresistible to the viewer. For example, a flying mosquito captures our attention while reading even though we try to ignore it and concentrate on reading.

Posner invented a spatial cueing paradigm that measures the effectiveness of attentional orienting by reaction time [42], in which the participants pressed a key instantly upon the discovery of a target appearing in the right or left visual hemifield. Before the presentation of the target, a flash cue was presented in one of the two hemifields. The target detection was found to be faster when the target was presented in the same hemifield as the flash cue. The result suggests that visual attention can be drawn to the cued location, and such attentional orienting helps target detection.

Both voluntary and involuntary attentional orienting can be effected by various spatial cues. Jonides used a central cue (an arrow pointing to a particular location) to elicit voluntary orienting and a peripheral cue (a flash appearing in the peripheral of the target location) to elicit involuntary orienting [25]. It was found that the central cue is susceptible to the participants' expectation, task difficulty, and cue validity (the probability that the cue and the target appear in the same location). In contrast, the peripheral cue can draw attention more automatically and is relatively more robust with respect to the cueing effect than the central cue. Müller et al. examined the temporal characteristics of voluntary and involuntary attention guidance [40] by delaying the presence of the target relative to the cue. The cueing effect was found to be maximal for the central cue when the delay was 300 ms. In contrast, the maximal cueing effect for the peripheral cue occurred when the delay was 175 ms. The result suggests that we may use a peripheral cue with instant cue-led time to elicit involuntary orienting [39]. This is preferred in our work because, as mentioned in Section 1, a visible cue irrelevant to the task at hand is distracting and may annoy the viewer and because attention guidance by invisible stimulus is more automatic and resistant to participants' intention than visible stimulus [8, 37].

McCormick found that visible and invisible cues demonstrate dissociated effects on attention guidance [56]. When the cue was visible (high contrast), participants voluntarily used the cue validity information and responded faster when the target was presented in the visual hemifield opposite to the cue (because where the probability of the target presence was higher) than in the same visual hemifield as the cue. In contrast, when the cue was invisible (low contrast), participants were involuntarily captured by the cue and it was found that participants responded faster when the target appeared in the same visual field as the cue (even though it had a lower probability of target presence) than in the opposite visual field. This interesting result suggests that a visible cue can guide voluntary attention, whereas an invisible cue can guide involuntary attention.

Common techniques used to create invisible stimuli include the following: limiting the presentation time, making a low contrast presentation, leveraging inter-ocular suppression, and forward/backward masking. For example, Tsuchiya et al. invented a method called "continuous flash suppression" that presented a visual stimulus to one eye and continuously flashes color masks to the other eye [22]. This discrepancy in visual input to the two eyes resulted in inter-ocular suppression where the strong dynamic masks for one eye masks the weak, low contrast visual stimulus from the other eye. As a result, the overriding mask is visible and the weak visual stimulus is invisible. Although this method is successful in making stimulus invisible even for complex visual stimulus [56, 62], it does not work in real-world applications since one cannot project discrepant visual inputs to two eyes in a natural environment. Further, the dynamic color masks may cause undesirable image quality degradation. Mulckhuyse et al. used limited presentation time and backward masking to make visual stimuli invisible [38, 39]. Location masks were presented immediately after the cue. It was found that the participants detected the target better when the target appeared at the cued location, suggesting that an invisible cue can guide attention to the cued location and that backward masking can make the cue invisible without causing undesirable distraction to the viewer. However, since only the left and the right visual hemifields were considered, it is unclear how the precision of this approach can be further enhanced.

3 System description

Our subliminal cueing system directs the visual attention of a viewer to a specific part of a real-world color image without engaging the viewer's awareness by briefly flashing a cueing image before displaying the real-world image. The block diagram of the system is depicted in Fig. 1a. Given the real-world image and the position that visual attention is to be directed to, this system automatically generates the cueing image (referred to as "cue" in the remainder of the paper) through a cue generator. The binary precision parameter is input to the cue generator to select the cue type. As shown in the flow chart in Fig. 1b, if low precision is selected, the system generates a blob cue centered at the cue location (Section 4.2). If high precision is selected, the system first detects if the cue location is inside an object and, if yes, generates an object cue (Section 4.4); otherwise, it generates a blob cue. A two-step process is adopted to generate a blob or an object cue. In the first step, we generate a blob image with black background. For a blob cue, the center of the blob would locate at the input position of the system. On the other hand, for an object cue, the center of the blob would locate at the center of the minimum circumscribed circle of the object which is at the input position of the system. After the blob image is generated, the cue generator blends it with the input real-world image

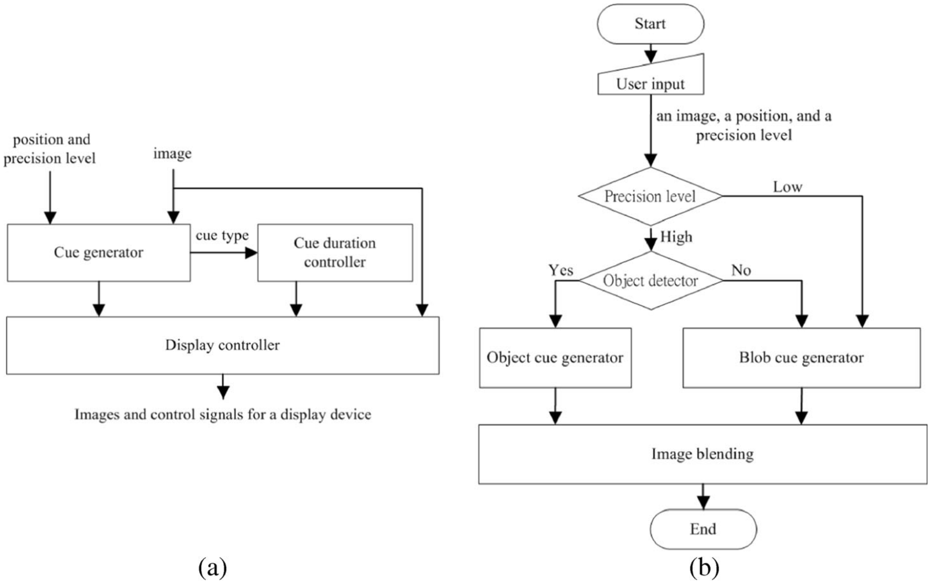


Fig. 1 **a** Block diagram of the subliminal cueing system, and **b** flow chart of the cue generator

to create the cueing image. The display duration of the cueing image is determined by a cue duration controller. Finally, the display controller sends the images and the display duration information to a display.

Although subliminal cueing has been studied intensely, which type of cue is more effective for practical applications is still an open issue. There is no common objective metric for evaluating the performance of a cue. To address this, we develop four experiments (Section 4) for subjective performance evaluation of the proposed blob cue and object cue and for investigating the precision of each kind of cue. Since human face is the most seen object in images, we separate the evaluation of face cue (Section 4.3) from that of other object cues.

All the approaches discussed in Section 2 use simple uniform gray-scale images for visual guidance. Whether these approaches work for color images of complex scenes in the real world should be addressed as well. For this purpose, these experiments are set up in a real-world environment, not in an idealized lab environment. The design considerations of these experiments are described in subsequent sections.

3.1 Competition for attention

As described earlier, psychology studies [22, 56] have shown that a short-duration cue displayed before an image can unconsciously guide the visual attention of a viewer to a certain targeted image location or area. In this context, the short-duration cue is an attention attractor. Likewise, image regions with high contrast in color, luminance, or orientation are attention attractors as well. Personal experience, which also affects visual attention, is another attention attractor. For example, most observers tend to pay attention to the middle part rather than the periphery of an image because experience has taught that attention attracting regions are usually located in the center of an image [51]. These attention attractors are segregated in most psychological studies through a proper control to ensure that only one attention attractor is in effect. In the real world, however, more than one attractor may exist at the same time and compete for attention.

Normally, a person can only attend to a limited number of stimuli at any time [61], and only strong attention attractors have an impact on the final deployment of visual attention. Therefore, we need to test the strength of the short-duration cue as an attention attractor against others. Since tracking eye movements is considered a reliable means for studying human visual attention [63], we use an eye tracker (Section 4.1) to record eye movements and evaluate the impact of short-duration cues on visual attention in the presence of other attention attractors. We try to cover as many competing attractors as possible in our experiments. Test images of various scenes and from various public websites, including Flickr [10], Free Stock Photos [16], Waterloo [15], Kodak [28], and SIPI [52], are used in these experiments as stimuli (Sections 4.2–4.4).

3.2 Visual cues

Visual attention is a selection process in the human brain to decide where or what to focus on within the visual field. Previous studies in psychology suggest two complementary modes of the selection process: space-based mode and object-based mode [7, 35, 47]. In the space-based mode, a location in the visual field is selected, whereas in the object-based mode, an organized chunk of visual information in the form of an object is selected. Accordingly, we use a blob cue (Section 4.2) to trigger space-based attention and a face or object cue (Sections 4.3–4.4) to trigger object-based attention. The blob is chosen because it is small enough to be placed at any location of an image. The face is chosen because it is often the object of interest in an image.

The threshold in subliminal stimuli research is the level at which the participant is not aware of the stimulus being presented [23], and the way to make the stimuli invisible depends on the type of the stimulus. For the blob cue used in this work, the duration of 50 ms is chosen for the following two reasons. First, according to the findings in previous studies [4, 5, 17, 46], a duration below 50 ms is usually used for a spatial cue with similar size and contrast as the blob cue. Second, the real-world complex image considered in our design would introduce backward masking effect, which allows us to select a cue duration close to the upper limit in the literature, i.e., 50 ms. For the face and object cues, since the cue is partly overlapping with the image, the effect of backward masking is excepted to be stronger than the case of blob cue. Since both of the cues are newly proposed, we select the duration of the cues based on subjective experiments (Sections 4.3–4.4), where the cue duration being tested are 50, 100, and 200 ms. The durations were empirically chosen by the authors.

3.3 Performance measures

Most psychological studies of short-duration cues primarily focus on their capability in directing human visual attention. For real-world applications, there is a need to measure the performance of a visual cue in terms of accuracy (how precisely a visual cue can draw human attention to the cued location) and efficiency (how fast a visual cue can draw human attention). These performance measures are used in our experiments and discussed in Sections 4.2–4.4.

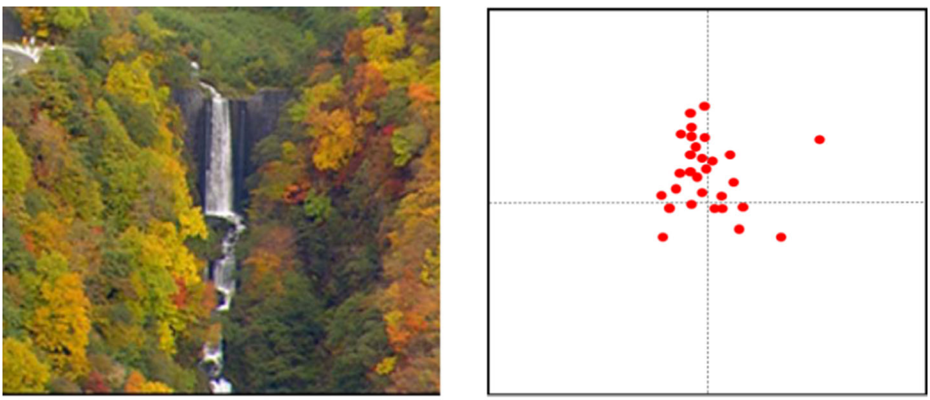
4 Evaluation

The performance of the subliminal cueing system described in Section 3 is evaluated using an eye tracker and images of real-world scene s. This section describes the experimental setup of four experiments and the eye tracker.

4.1 Eye tracking apparatus

As visual attention and eye movements are correlated [12, 19, 54], we use an eye tracking device to locate the image regions of interest in our experiments. An eye tracker can record various eye movement events, including fixation, saccade, and blink. The eyes initially fixate on a specific location then jump to another location, fixate for a while, and jump to another location. They undergo a jump-and-rest scan pattern. Each jump is called a saccade, and there are about three saccades per second. When the visual attention is attracted to a particular area or object in the image, most saccades point toward it. The degree of attention varies from region to region. The higher the fixation density, the more attractive a region is. An example is shown in Fig. 2, where the waterfall is the most attended region since it has the highest fixation density relative to the other regions (e.g., trees) in the image.

The eye tracking system (EyeLink 2000, SR Research Ltd) used in our experiments consists of an infrared camera, a host PC, a display PC, and a chin-and-forehead rest as shown in Fig. 3. The infrared camera tracks the eye movements of the viewer, the host PC records the eye movement data, and the display PC presents the stimuli to the viewers. The system uses monocular recording, with a resolution of 0.02° RMS. The eye movement-recording rate is 1000 Hz, and the refresh rate of the display is 60 Hz. At the beginning of each experiment, a calibration procedure with a nine-point pattern as shown in Fig. 3c is performed to register the eye position with the pixel location on the monitor.



(a) (b)

Fig. 2 An example of human fixation data: **a** The original image and **b** the fixation points collected from 16 subjects

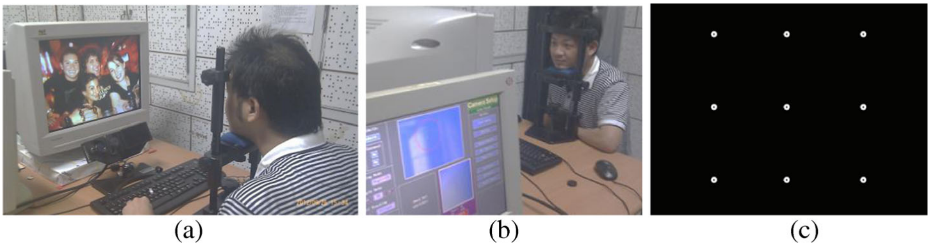


Fig. 3 The eye tracking system used in our experiments: **a** The setup of the eye tracking experiment, **b** The chin-and-forehead rest, and **c** The nine-point calibration pattern

4.2 Experiment 1 (blob cue)

This experiment examined whether a short-duration spatial cue (a blob) presented before the display of a real-world image attracts an observer's attention. We invited twelve students from 18 to 22 years old at our university to participate in this experiment. All had normal or corrected-to-normal vision and were naïve about the purpose of this experiment.

The stimuli were displayed on a 22-in. ViewSonic P225f CRT monitor (1024×768 resolution with 60 Hz refresh rate) controlled by an IBM compatible personal computer with the Psychophysics Toolbox of Matlab software. In the experiment, which consists of a number of trials, the tracking system described in Section 4.1 recorded eye movements of each subject and a gray blob (50 pixels in radius) with Weber contrast value of 10 % was used as the spatial cue. In each trial, the spatial cue was presented for 50 ms at a randomly assigned location, followed by the presentation of an image (1024×768 pixels) for 3000 ms. A set of 240 images obtained from the Kodak image database and a number of other public websites were used as test images. These images were originally taken from various natural scenes, cities, creatures, and other scenes.

The procedure of this experiment is depicted in Fig. 4. Each participant sat at a viewing distance 60 cm from the display and fixed his or her head on a chin-and-head rest. A drift correction was conducted prior to each trial, in which each participant had to fixate on a central white point of 5×5 pixels and concurrently pressed the space bar on the keyboard to initialize the trial. The purpose of this calibration step was to make sure that the eyes fixated at the same location at the beginning of each trial to avoid bias towards particular visual field. This step was also necessary for obtaining accurate eye movement data over trials. After the trial started, participants were allowed to view the images freely. Note that we made the gray blob very low contrast (10 % in Weber's contrast) and its presentation time very brief (50 ms).

A characteristic of HVS is that the left and right hemifields have independent attentional object tracking mechanisms [2]. Therefore, if the short-duration blob cue can attract participants' visual attention, the fixation is very likely to be directed toward the cued left or right hemifield. To probe if such attentional guidance exists, we calculate P_e, the probability of "eye on cued visual field", for the first and the second fixations. If the "location of cue" is independent of the "location of fixation", P_e should be 0.5 because the blob cue is randomly presented in left and right visual fields. If P_e is higher than 0.5, it means that the location of cue has an effect on the location of fixation, and a larger value of (P_e-0.5) indicates a stronger effect of cue location that attracted fixation. The results shown in Fig. 5 indicate that the second fixation is significantly biased toward the cued hemifield, ($t(11)=3.93$, $p<0.01$). Here 3.93 is the t-test coefficient (t-value), and the degree of freedom is 11 because there are 12

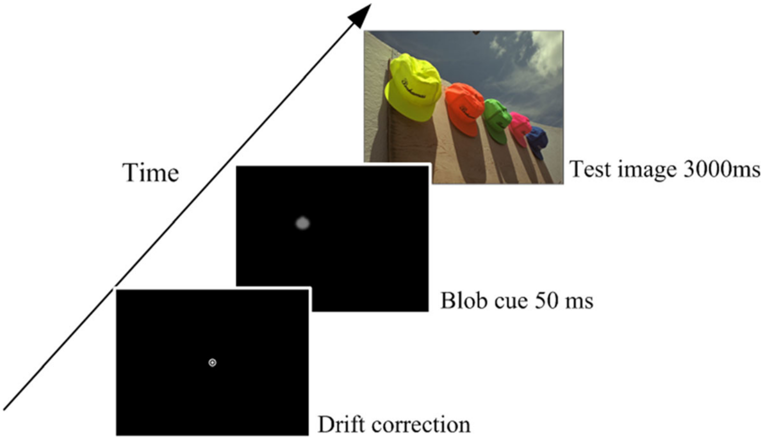


Fig. 4 The procedure of Experiment 1. Each trial begins with a drift correction, in which the participant has to fixate on a central white point and concurrently press the space bar on the keyboard to continue the trial. Next, a randomly located *gray blob* is presented for 50 ms. Finally, a test image is presented for 3000 ms

participants in this experiment. The proportion of the t-distribution beyond the t-value (p-value) is smaller than 0.01, which means that the probability of the sampled second fixation points which comes from the same sampling distribution with the mean as 0.5 (i.e., the chance level) is smaller than 0.01. Therefore, it is concluded that at a 99 % confidence level [59], the second fixation is dependent on the location of cue. However, no such effect was found for the first fixation.

To further examine whether the short-duration spatial cue attracts attention to the exact location, we computed the probability that the first and second fixations were in the cued quadrant. The chance level now became 0.25 because there were four different cued quadrants. The results shown in Fig. 6 indicate that neither the first nor the second fixation is biased toward the cued quadrant (for the first fixation, $t(11)=0.88$, $p<0.3$, and for the second fixation, $t(11)=0.96$, $p<0.3$). Together with the above cued hemifield analysis, we conclude that a low contrast blob cue presented for a short duration is able to attract visual attention to the cued

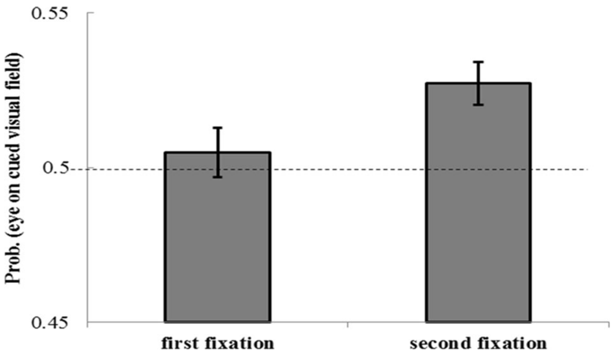


Fig. 5 Mean probability distribution of the first and second fixations being directed to the cued hemifield in Experiment 1. An *error bar* indicates the standard deviation of the probability distribution of each fixation. The results show that the second fixation is biased towards the cued hemifield, but no such bias is found for the first fixation

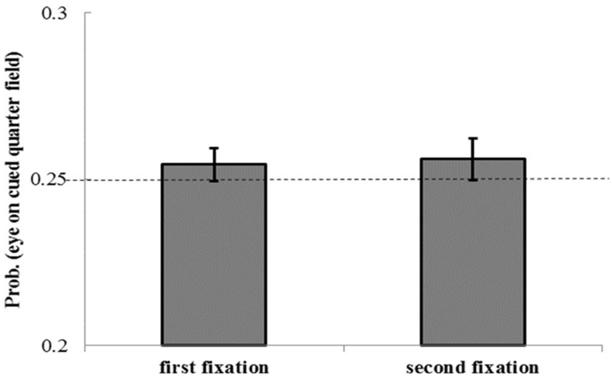


Fig. 6 Mean probability distribution of the first and second fixations being directed to the cued quadrant in Experiment 1. The results show that neither the first nor the second fixation is biased toward the cued quadrant. An *error bar* indicates the standard deviation of the probability distribution of each fixation

hemifield, but the precision is not enough to direct visual attention to the cued quadrant (or cued location).

4.3 Experiment 2 (face cue)

The lack of cueing precision for a blob cue may be due to the fact that the cue is only a simple geometric shape. The purpose of this experiment is to verify that an object cue such as a face leads to a higher cueing precision.

Another group of 24 students aging from 18 to 22 at our university was invited to participate in this experiment. The apparatus and design of this experiment were the same as Experiment 1 described in Section 4.2 except for the following. First, we used 20 images from Flickr with a Creative Commons license [10], each image (1024×768 pixels) containing four faces. Each face was cued with equal probability and for three different cue durations, 50, 100, and 200 ms. A between-subjects design [60] that randomly and equally divided the participants into three groups was used, each group with a different cue duration. The test images were displayed in random order to remove the effect of presentation order on visual attention. Second, the face cue was generated from the face image to be displayed so that the difference between the cue and the displayed image was small and hardly to be noticed during the short cuing duration. To make an object cue from a face image, we first applied the face detector developed by Nilsson et al. [41] and circumscribed each detected face by a circle. Then we gradually attenuated the luminance of the cued face region from the center of the circle to the boundary. The luminance of the remaining region was also attenuated to avoid any visible edge at the boundary of the face region. Let $I(x, y)$ denote the displayed image and (x_f, y_f) and r_f, respectively, denote the center coordinates and the radius of the cued face region. The cue image $C(x, y)$ is obtained by

$$C(x,y)\begin{cases} I(x,y)e^{-r^2/r_f^2}, & r^2 < r_f^2 \\ I(x,y)\, e^{-1}, & \text{otherwise} \end{cases}. \tag{1}$$

where x and y, respectively, denote the horizontal and vertical coordinates of a pixel, and

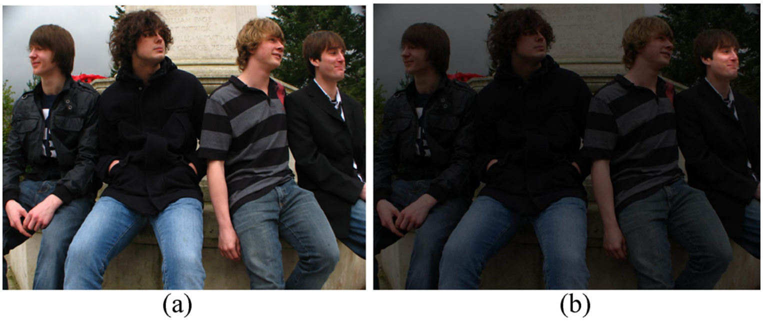


Fig. 7 **a** An image and **b** its corresponding face cue. In this example, the face of the first person from right is selected as the cue

$r^2=(x-x_f)^2+(y-y_f)^2$. An example image and its corresponding face cue thus generated are shown in Fig. 7.

To examine whether the face cue can guide visual attention to the cued hemifield, we computed the probability of the first and the second fixation points in the cued hemifield and compared them with the chance level 0.5. The results shown in Fig. 8 indicate that the second fixation is biased toward the cued hemifield when the cue duration is 200 ms ($t(7)=3.33$, $p<0.02$), but such gaze bias is not found in the other cases. To further examine whether the face cue has a higher precision in location, we computed the conditional probabilities of the first and second fixations being guided (biased) by the cued face given that each of the first and second fixations was on a face. If a fixation point was indeed directed to the cued face, the conditional probability should be greater than the chance level 0.25. We can see from the results in Fig. 9 that the second fixation is biased towards the cued face when the cue duration is 200 ms ($t(7)=3.57$, $p<0.01$), but no such gaze bias was found in the other cases. This analysis, in conjunction with the cued hemifield analysis in the experiment described in Section 4.2, indicates that directing the visual attention to a cued location (and hence the cued hemifield) can be effected by using a face cue with 200 ms cue duration.

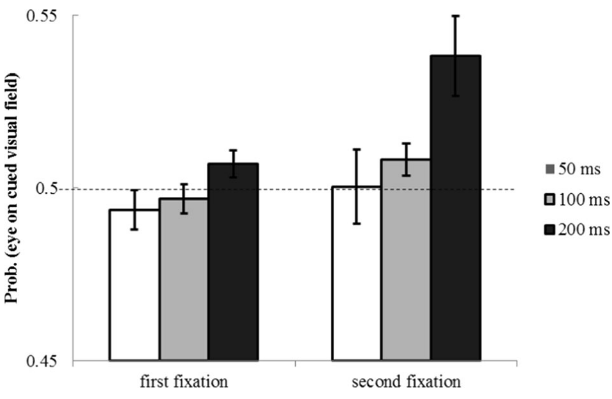


Fig. 8 Mean probability distribution of the first and second fixations being directed to the cued hemifield in Experiment 2. The results show that the second fixation is biased towards the cued hemifield when the cue duration is 200 ms. An *error bar* indicates the standard deviation of each probability distribution

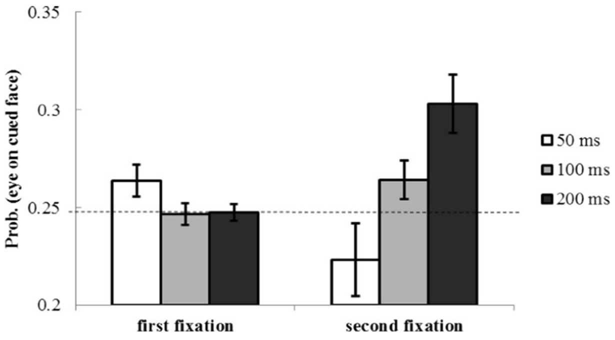


Fig. 9 Mean probability distribution of the first and second fixations being directed to the cued face. The results show that second fixation is biased towards the cued face when the cue duration is 200 ms. An *error bar* indicates the standard deviation of each probability distribution i

To understand more about how the cue duration affects attentional guidance, we performed an analysis of variance (ANOVA) test [58] at 99.5 % confidence level [60] to determine the reliability of the conditional probability differences between different cue durations for the second fixation. A highly significant simple main effect [48] was found. That is, the face cue with longer duration led to a more effective attentional guidance. This also suggests that it takes time for the face cue to bring attentional guidance into effect and that to be most effective, the cue duration should be as close as possible to the subliminal threshold.

To quantify how the longer cue duration affects attention guidance, we computed the increase of fixation points inside the cued face region when the cue duration increased from 50 to 200 ms. The results listed in Table 1 show that the increment rate is 80 % for the first fixation, 143 % for the second fixation, and 130 % for all the fixations within the 3000 ms presentation duration of the test images. The results show that the attentional guidance of the face cue can last for at least 3000 ms and that the attentional guidance is strongest at the second fixation and decreases over time. Figures 10 and 11 show the fixation data of 50 and 200 ms cue durations for one test image. As we can see, the attention guidance becomes stronger as the cue duration increases. Moreover, since the attention guidance of the 50 ms cue is not statistically significant as described earlier in this section, the attention guidance power of the 200 ms cue (which scores 730 in the second fixation) can be considered to be 243 % of the uncued case.

To find out if a short-duration face cue has an additional effect on the distribution of focal points, we compared the fixation density map (which is the probability density map of the fixation points) of the cued face regions with that of the uncued face regions. Since the cued face regions were at different locations and had different sizes from trial to trial, a normalization of the fixation data was needed. Considering that each cued face region was described by a

Table 1 The total number of fixation points inside of the cued face

	1st fixation	2nd fixation	Fixations within 3000 ms
50 ms	153	300	2239
200 ms	276	730	5150
Ratio[a]	1.80	2.43	2.3

[a] The ratio of the number of fixation points inside of the cued face between the two cue durations

(a) (b)

Fig. 10 Fixation data for a test image: **a** 50 ms cue duration and **b** 200 ms cue duration. The *red dots* are the fixation points inside the cued face (the third face from left) in the second fixation. This example illustrates that the attention guidance becomes stronger as the cue duration increased from 50 ms to 200 ms

circumscribing circle centered at (x_f, y_f) with radius r_f, we mapped each fixation point (x_p, y_p) inside the circle to a point (x_n, y_n) inside a normalized circle centered at (0, 0) and with radius r_m, the average radius of the cued face regions, by the following formula:

$$(x_n, y_n) = \left[\left(x_p - x_f\right) \frac{r_m}{r_f}, \left(y_p - y_f\right) \frac{r_m}{r_f} \right]. \quad (2)$$

In other words, the centers of the circles were aligned first. Then the distance of each fixation point to the center was proportionally scaled. Fixation data was convolved with a Gaussian filter to produce the final fixation density map of the cued face regions [29]. The same procedure was applied to generate the fixation density map of the uncued face regions. The results are shown in Fig. 12. A Student's T test [55] of the normalized fixation data yielded a p-value of 0.95, which indicates that the distribution of the fixation data is statistically indifferent between the cued and uncued images. In other words, the face cue brings the viewer's attention to the cued face but it does not affect the distribution of attention within the face region.

(a) (b)

Fig. 11 Fixation data for a test image: **a** 50 ms cue duration and **b** 200 ms cue duration. The *red dots* are the fixation points inside the cued face (the fourth face from left) in the second fixation. This example illustrates that the attention guidance becomes stronger as the cue duration increased from 50 ms to 200 ms

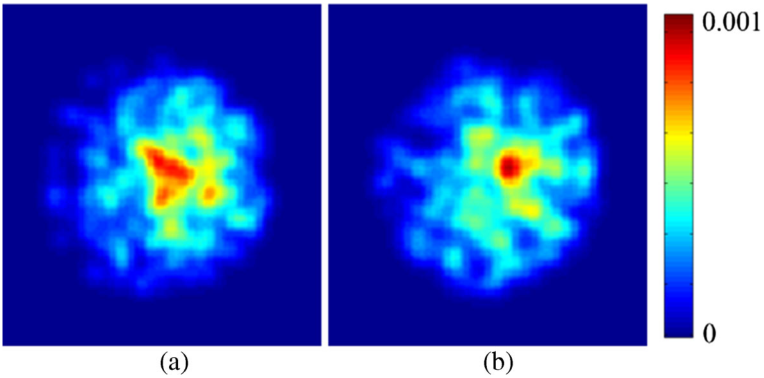


Fig. 12 Normalized distribution of fixation points of the second fixation for **a** cued face region and **b** uncued face region

4.4 Experiment 3 (object cue)

The goal of this experiment was to examine if a generalized object cue can subliminally attract human visual attention. This experiment can tell us the range of applications for object cues. A positive result means that the proposed attention guiding method may be effective for images of any objects. If the result is negative, the proposed method is only effective for images of face. The details of the experiment are described in this section.

4.4.1 Stimuli selection

Theoretically, the proposed attention guiding method should be tested on all kinds of object cues for performance evaluation. In practice, however, this would result in an extremely long experiment and is also an extremely large subjective load on participants. To efficiently evaluate the performance of object cues, we adopted a divide-and-conquer method by randomly sampling object cues from seven representative object categories. If these seven object categories were selected in a psychologically meaningful manner, then our findings from the sampled object cues can be generalized to different kinds of object cues.

An object categorization method based on brain activity analysis [1] is adopted in this experiment for picking out representative object cues. This method is chosen because it categorizes objects based on aspects of psychological signals instead of synonyms because the latter may not be strongly correlated with human visual attention. In detail, in the categorization method, Alexander et al. [1] analyzed the relation between fMRI measurement of the Blood-oxygen-level dependent (BOLD) and humans' ability to name thousands of distinct objects through principal components analysis (PCA). Their results show that with the first to the fourth principle components (PCs), the fMRI measurement can be effectively related to more than 49 % of the testing objects. Moreover, the four PCs have semantic interpretations. The first PC distinguishes between moving and still objects, the second PC distinguishes between indoor social interaction and outdoor activities, the third PC distinguishes between nature and civilization, and the fourth PC distinguishes between living and non-living objects.

The second to the fourth PCs were adopted in this experiment for stimuli selection. The first PC was not adopted because the stimuli used in this experiment were mainly still images, and it is hard to tell if an object is going to move from a still image. With the second to the fourth PCs, we categorized the testing objects used in [1] into 8 groups, each of which represented a combination of the three PCs. The results are shown in Table 2. Since there is no object corresponding to (−, +, +) from the 1705 objects used in [1], the stimuli used in this experiment were eventually selected from the rest of the groups (person, plant and animal, structure, craft, text, vehicle, furniture).

4.4.2 Experiment design and results

Another group of 17 students aging from 18 to 22 at our university was invited to participate in this experiment. The apparatus and design of this experiment were the same as those of the experiment described in Section 4.3 except that all of the object cues were presented for 100 ms.

Images from the seven object categories mentioned above were used as the stimuli. For each category, four test images that contained at least two objects that belonged to it were selected. All of the images were picked from Scene Understanding database (SUN) [49], which contains a comprehensive collection of annotated images covering a large variety of environmental scenes, places and the objects.

To make an object region into an object cue, we decreased the luminance of the image except for this object region. Specifically, we first detected the object regions in the image, and each of the detected objects was described by its circumscribing circle. Then we gradually attenuated the luminance of the cued object region from the center toward the boundary. We also attenuated the luminance of the other regions to avoid visible edges in the boundary of the object region. Let I denote the displayed image, (x_f, y_f) and r_f, respectively, denote the central coordinates and the radius of the circle corresponding to the cued object region.

The formula for generating a cue image C is as follows:

$$\begin{aligned} C(x,y) &= I(x,y)e^{-r^2/r_f^2}, \quad r^2 < r_f^2, \\ C(x,y) &= I(x,y)\,e^{-1}, \quad \text{otherwise}. \end{aligned} \tag{3}$$

where x and y, respectively, denote the horizontal and vertical positions of a pixel, and $r^2=(x-x_f)^2+(y-y_f)^2$.

Table 2 Object categorization based on PCA analysis

(2nd PC, 3rd PC, 4th PC)	Object category
(+, +, +)	Text
(+, +, −)	Plant and animal
(+, −, +)	Craft
(+, −, −)	Vehicle
(−, −, +)	Structure
(−, +, −)	Person
(−, −, −)	Furniture

To explore whether the object cue drew participants' attention to a specific object, the conditional probability was computed for the following three conditions:

- Cue: The first or second fixations was located at the cued object given that an object cue was used
- Un-cue: The first or second fixations was located at another object that is the same category as the cued object given that an object cue was used
- No-cue: The first or second fixations was located at the cued object given that no object cue was used

The results are shown in Figs. 13 and 14. According to ANOVA statistics, there were statistically significant differences between "cue" and "no cue" situation and between "cue" and "un-cue" situations with the second saccade. However, there were no statistically significant differences between "un-cue" and "no cue" situation. The results indicate that the object cue can direct human attention to the exact object we cued. Other objects in same categories should not be affected.

4.5 Experiment 4 (subliminality test)

This experiment was conducted to examine whether the short duration cues used in the previous experiments were subliminal. We invited another 18 students aging from 18 to 22 at our university to participate in this experiment. All had normal or corrected-to-normal vision and were naïve about the purpose of this experiment.

Participants' visual attention allocation can be different from their natural viewing situation when they were aware of the goal of the experiment. For example, if we explicitly ask the participants if they see anything presented before an image, it is likely that they would concentrate on finding something presented before the image instead of viewing the image. Therefore, the procedure of this experiment is carefully designed to avoid affecting the attention allocation of the participants.

This experiment had 3 sessions. One for the 50-ms blob cue, another one for the 200-ms face cue, and the third one for the 100-ms object cue. All sessions contained 20 trials. In each trial, we showed the participant a cue embedded image and asked them to tell us their first impression of the image. There were two options for them: positive or negative. The purpose of this question was to make the participants concentrate on the image viewing task. At the end

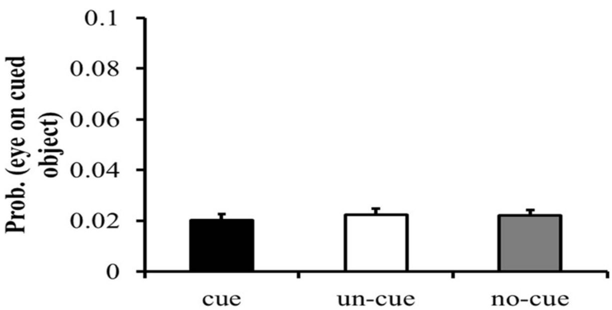


Fig. 13 Conditional probabilities of the cue, un-cue, and no-cue objects for the first fixation. The results show that the first fixation is not biased toward any object in the displayed image

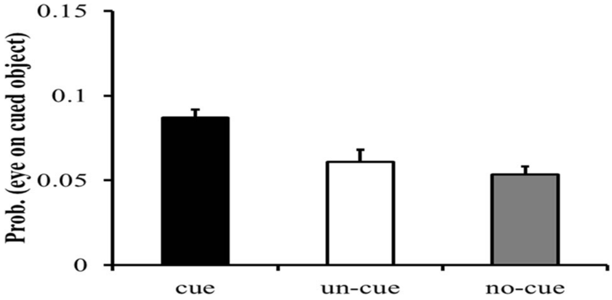


Fig. 14 Conditional probabilities of the cue, un-cue, and no-cue objects for the second fixation. The conditional probabilities show that cue object is more likely to be seen, meaning that the second fixation is biased towards the cue object

of the experiment, we asked the participants to tell us if they saw anything abnormal between the trials and asked them to rate how confident they were in a 5-point Likert scale [33]. The purpose of this question was to test if the cues were detected.

The experimental results are shown in Table 3, where the detection rate refers to the percentage of the participants' seeing the cue, and the average confidence level refers to the average score in the 5-point Likert scale. The results show that none of the participants detected the 50-ms blob cue and the 100-ms object cue, and only 5.6 % (1 out of 18) of the participants detected the 200-ms face cue. The only participant who detected the cue reported that she saw a flash before the image but she was not sure what was flashed. The average confidence levels are both higher than 4, which indicated that the participants were confident about their answers.

5 Discussion

It has been shown that short-duration subliminal cues can attract human visual attention in simple displays in the laboratory such as in random-dot images or images consisting of simple geometric figures [36], [43], however it remains unknown about the effect of subliminal cues in real-world images. Experiment 1 showed that a blob cue presented for 50 ms attracted human gaze to the cued hemifield but not to the cued quadrant. Experiments 2 and 3 showed that a 200-ms face cue and a 100-ms object cue attracted human gaze to the cued location. Experiment 4 confirmed that the short-duration cues which direct human visual attention in the real-world image did not engage the viewer's awareness. A statistical analysis of the results of all experiments showed that the gaze bias predominantly occurred in the second fixation but not the first one.

These results have two important implications. First, that three kinds of short-duration cues (blob, face, and object) can direct human visual attention suggests that visual attention can be deployed before the perception of scene content is formed. Second, the attentional guidance effect of the short-duration visual cues is not negligible even when they do not engage human awareness. That is, despite the fact that in the test images there are other high contrast regions, including faces, which are attention attractors as well [29], the short-duration cues used in our experiments are able to affect attentional orienting.

The results of the four experiments are helpful in the understanding of the processing sequence of human visual attention. Recall that a 50-ms duration is sufficient for a blob cue to

Table 3 The detection rate of the subliminal cues

	50-ms blob cue	200-ms face cue	100-ms object cue
Detection rate	0 %	5.6 %	0 %
Average confidence level	4.05	4	4.36

attract human attention, a 100-ms duration is sufficient for an object cue, whereas a 200-ms duration is needed for a face cue. The longer cue duration required by the face and object cue implies that HVS needs more time to process a brightening cue than an added cue, and this can all happen before the deployment of spatial attention to the cue. Lo et al. observed a similar phenomenon [34]. Their results show that it takes time for HVS to process a still image even before spatial attention is deployed to it. Our experiments extend their observation to flashed images. Also, our results show that the extra processing time for the face cue facilitates the direction of attention to a more specific location—the cued face.

The results of Experiments 2 and 3 showed that different durations were needed for different types of cues: 100 ms was sufficient for object cues but 200 ms was necessary for face cues. As a human face has higher ecological meaning and is processed more automatically than a nature scene, it is interesting to discuss the reason why faces require more time for attentional orienting than other objects. One possibility is due to the content of images we used in different experiments. There were always four faces in the face cue condition, and it is possible that some un-cued faces also attracted attention [26] and thus diluted the cueing effect. Therefore, longer time was needed for each face to attract attention.

Someone might argue that the durations we used for each type of cues were too long to be subliminal. Indeed, in our study, the duration is much longer than what has been found in the simple display situation [38]. However, the cue that was used in their study had sharp boundaries (solid circle on a white background) and was always presented in the same location, which would induce high contrast and predictable spatial information. In this work, we used the blob with smooth boundary and it was randomly presented in the whole visual field. In addition, our method introduces an interaction between the cue and the color image of the complex natural scene. This interaction makes the color image a backward masking pattern for the cue and prolongs the subliminal processing.

It is interesting to find that the gaze bias as an effect of subliminal cuing predominantly occurs in the second fixation, not the first one. A possible reason of the phenomenon is that there may be some influence of the calibration procedure on the first fixation, or that the human tends to first look at the center of a displayed image, which is known as the central bias tendency [26]. If such a tendency is strong, it confines the visual attention of the first fixation to the central area of an image. Thus, unless the cued location happens to be near the image center, our eyes will not fixate on it in the first fixation. To examine whether the central bias tendency exists in our experiments, we plotted the fixation data of all test images. The results confirmed that the central tendency does exist and that it is stronger in the first fixation than the second and the third fixations. For illustration purposes, the results for two example images are shown in Fig. 15. In each example, the left image

(a)

(b)

Fig. 15 Illustration of the central tendency of visual attention to displayed images. The fixation points of the first fixation of all subjects are shown on the left image, and those of the second and third fixations are shown on the right image. We can see that most fixation points of the first fixation are located in the central region of the image—an evidence of the central bias tendancy

contains the fixation points of the first fixation, and the right image contains the fixation points of the second and the third fixations. We can see clearly that the distribution of fixation points is indeed in accordance with the central bias tendency.

From the results of our experiments, it is reasonable to expect that the proposed object-based cuing method can be applied to a wide range of objects of interest in practice. The results of Experiment 3 are directly supportive of this statement, and the results of Experiments 1 and 2 can be viewed as the performance of object cues for corner cases. Specifically, the blob used in Experiment 1 can be considered an object with little details and the face used in Experiment 2 can be considered an object with great details. It is reasonable to expect that objects with richness of detailed information in between a blob and a face should be able to direct human visual attention as well, and the valid cue duration is expected to be between 50 and 200 ms.

Further extensions of this work can be pursued. One possibility is to examine, from the attention guidance point of view, the interaction between the proposed short-duration cues and other attention attractors described in Section 3. Another possible direction of future work is the generalization of the short-duration cuing method to video sequences. It is anticipated that issues such as the arrangement of cued frames and the balance between viewing quality and attention guidance performance are worth studying further.

6 Conclusion

We have described a subliminal cuing method that directs human visual attention to real-world images. Our method contains three types of cues, which have different properties. One is a spatial cue that could be used with all kinds of images by just flashing a dim blob for 50 ms. The cue is able to direct attention to the cued hemifield. The other two types of cues require face/object detectors processing of the image, and they can direct attention to the precise cued location. However, the cues require more time to guide attention—that is, 200 ms for face cue, and 100 ms for object cue. Since the cues we discussed here can attract viewer's attention unconsciously, the non-intrusive attentional guidance technique described in this paper is useful for many multimedia applications such as digital signage, advertisement media design, digital art, and education.

References

1. Alexander G, Huth L, Nishimoto S, An T, Jack LG (2012) A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron 76(6):1210–1224
2. Alvarez GA, Cavanagh P (2005) Independent resources for attentional tracking in the left and right visual hemifields. Psychol Sci 16(8):637–643
3. Blanchfied A, Hardy J, Samuele M (2014) Non-conscious visual cues related to affect and action alter perception of effort and endurance performance. Front Hum Neurosci 8(1):1–16
4. Burnette KE, d'Avossa G, Sapir A (2013) Matching cue size and task properties in exogenous attention. Q J Exp Psychol 66(12):2363–2375
5. Busse L, Katzner S, Treue S (2006) Spatial and feature-based effects of exogenous cueing on visual motion processing. Vis Res 46(13):2019–2027
6. Chalfoun P, Frasson C (2011) Subliminal cues while teaching: HCI technique for enhanced learning. Advances in Human-Computer Interaction, article ID 968753 1–15
7. Chou WL, Yeh SL (2008) Location- and object-based inhibition of return are affected by different kinds of working memory. Q J Exp Psychol 61:1761–1768
8. Chou WL, Yeh SL (2011) Subliminal spatial cues capture attention and strengthen between-object link. Conscious Cogn 20:1265–1271
9. Chou WL, Yeh SL (2012) Object-based attention occurs regardless of object awareness. Psychon Bull Rev 19(2):225–231
10. Creative Commons (2012) Flicker: creative commons. http://www.flickr.com/creativecommons/
11. Culibrk D, Mirkovic M, Zlokolica V, Pokric M, Crnojevic V, Kukolj D (2001) Salient motion features for video quality assessment. IEEE Trans Image Process 20(4):948–958
12. Duc AH, Bays P, Husain M (2008) Eye movements as a probe of attention. Prog Brain Res 171:403–411
13. Duncan J (1984) Selective attention and the organization of visual information. J Exp Psychol Gen 113:501–517
14. Egly R, Driver J, Rafal RD (1994) Shifting visual attention between objects and locations: evidence from normal and parietal lesion subjects. J Exp Psychol Gen 123:161–176
15. Fractal coding and analysis group (2009) Repository. http://links.uwaterloo.ca/Repository.html
16. Free Stock Photos.com (2011) Free Stock Photos.com. http://freestockphotos.com/
17. Fuller S, Park Y, Carrasco M (2009) Cue contrast modulates the effects of exogenous attention on appearance. Vis Res 49(14):1825–1837
18. Gibson BS, Bryant TA (2005) Variation in cue duration reveals top-down modulation of involuntary orienting to uninformative symbolic cues. Percept Psychophys 67(5):749–758
19. Groner R, Groner MT (1989) Attention and eye movement control: an overview. Eur Arch Psychiatry Neurol Sci 239(1):9–16
20. Huang TH, Yang YH, Liao HI, Yeh SL, Chen HH (2012) Directing visual attention by subliminal cues. IEEE International Conference on Image Processing: 1081-1084

21. Itti L, Dhavale N, Pighin F (2003) Realistic avatar eye and head animation using a neurobiological model of visual attention. Proc SPIE 48th Annu Int Symp Opt Sci Technol 5200:64–78
22. Jiang Y, Costello P, Fang F, Huang M, He S (2006) A gender- and sexual orientation-dependent spatial attentional effect of invisible images. Proc Natl Acad Sci USA 103(45):17048–17052
23. Jim C, Philip MM (1984) Priming with and without awareness. Percept Psychophys 36(4):387–395
24. Jin SM, Lee IB, Han JM, Seo JM, Park KS (2008) Context-based pixelization model for the artificial retina using saliency map and skin color detection algorithm. Proc. SPIE Human Vision and Electronic Imaging XIII 6806
25. Jonides J (1981) Voluntary versus automatic control over the mind's eye's movement. In: Laong JB, Baddeley AD (eds) Attention and performance, 4th edn. Erlbaum, Hillsdale, pp 187–203
26. Judd T, Ehinger K, Durand F, Torralba A (2009) Learning to predict where humans look. Proc. IEEE 12th International Conference on Computer Vision 2106–2133
27. Karremans JC, Stroebe W, Claus J (2006) Beyond vicary's fantasies: the impact of subliminal priming and brand choice. J Exp Soc Psychol 42:792–798
28. Kodak Lossless True Color Image Suite (1999) True color Kodak images. http://r0k.us/graphics/kodak/
29. Lee WF, Huang TH, Yeh SL, Chen HH (2011) Learning based prediction of visual attention for video signals. IEEE Trans Image Process 20(11):3028–3038
30. Li H, Ngan KN (2008) Saliency model-based face segmentation and tracking in head-and-shoulder video sequences. J Vis Commun Image Represent 19(5):320–333
31. Liao HI, Yeh SL (2011) Interaction between stimulus-driven orienting and top-down modulation in attentional capture. Acta Psychol 138:52–59
32. Libet B (1985) Unconscious cerebral initiative and the role of conscious will in voluntary action. Behav Brain Sci 8:529–566
33. Likert R (1932) A technique for the measurement of attitudes. Arch Psychol 26(140):1–55
34. Lo SY, Yeh SL (2008) Dissociation of processing time and awareness by the inattentional blindness paradigm. Conscious Cogn 17:1169–1180
35. Maioli C, Benaglio I, Sosta S, Siri K, Cappa S (2001) The integration of parallel and serial processing mechanisms in visual search: evidence from eye movement recordings. Eur J Neurosci 13:364–372
36. Mcauliffe J, Prat J (2005) The role of temporal and spatial factors in the covert orienting of visual attention tasks. Psychological Research 69:285–291
37. McCormick PA (1997) Orienting attention without awareness. J Exp Psychol Hum Percept Perform 23(1): 168–180
38. Mulckhuyse M, Talsma D, Theeuwes J (2007) Grabbing attention without knowing: automatic capture of attention by subliminal spatial cues. J VisCogn 15:779–788
39. Mulckhuyse M, Theeuwes J (2010) Unconscious attentional orienting to exogenous cues: a review of the literature. Acta Psychol 134(3):299–309
40. Müller HJ, Rabbitt PM (1989) Reflexive and voluntary orienting of visual attention: time course of activation and resistance to interruption. J Exp Psychol Hum Percept Perform 15(2):315–330
41. Nilsson M, Nordberg J, Claesson I (2007) Face detection using local SMQT features and split up snow classifier. Proc ICASSP 2:589–592
42. Posner MI (1980) Orienting of attention. Q J Exp Psychol 32:3–25
43. Pratt J, Hillis J, Gold JM (2001) The effect of the physical characteristics of cues and targets on facilitation and inhibition. Psychon Bull Rev 8(3):489–495
44. Reif F (2008) Applying cognitive science to education. The MIT Press, Cambridge
45. Ritter W (2011) Benefits of subliminal feedback loops in human-computer interaction. Adv Hum Comput Interact 2011(1):1–11
46. Sebastiani M, Casagrande M, Martella D, Raffone A (2009) The effects of endogenous and exogenous spatial cueing in a sustained attention task. Cogn Process 10(2):302–304
47. Soto D, Blanco MJ (2004) Spatial attention and object-based attention: a comparison within a single task. Vis Res 44:69–81
48. Stockburger DW Two way ANOVA and interactions. http://www.psychstat.missouristate.edu/multibook/mlt09m.html
49. Sun Database (2012) Scene understanding. http://groups.csail.mit.edu/vision/SUN/
50. Tang CW, Chen CH, Yu YH, Tsai CJ (2006) Visual sensitivity guided bit allocation for video coding. IEEE Trans Multimedia 8(1):11–18
51. Tatler BW (2007) The central fixation bias in scene viewing: selecting an optimal viewing position independently of motor biases and image feature distributions. J Vis 7(14):1–17

52. The USC-SIPI Image Database (1999) SIPI image database. http://sipi.usc.edu/database/. Accessed 15 Nov 1999
53. Theeuwes J (2010) Top-down and bottom-up control of visual selection. Acta Psychol 123:77–99
54. Torralba A, Oliva A, Castelhano MS, Henderson JM (2006) Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search. Psychol Rev 113(4):766–786
55. Trochim MK (2006) The T-test. http://www.socialresearchmethods.net/kb/stat_t.php. Accessed 20 Oct 2006
56. Tsuchiya N, Koch C (2005) Continuous flash suppression reduces negative afterimages. Nat Neurosci 8(8): 1096–1101
57. Wang YS, Tai CL, Sorkine O, Lee TY (2008) Optimized scale-and-stretch for image resizing. ACM Trans Graph 27(5):1–8
58. Wikimedia Foundation (2012). Analysis of variance. http://en.wikipedia.org/wiki/Analysis_of_variance. Accessed 5 Dec 2012
59. Wikimedia Foundation (2012) Confidence interval. http://en.wikipedia.org/wiki/Confidence_interval. Accessed 28 Nov 2012
60. Wikimedia Foundation (2013) Between-group Design. http://en.wikipedia.org/wiki/Between-group_design. Accessed 23 Apr 2013
61. Wolfe JM, Horowitz TS (2004) What attributes guide the deployment of visual attention and how do they do it. Nat Rev Neurosci 5:1–7
62. Yang YH, Yeh SL (2011) Accessing the meaning of invisible words. Conscious Cogn 20:223–233
63. Yarbus AL (1967) Eye movements and vision. Plenum, New York
64. Yeh SL, He S, Cavanagh P (2012) Semantic priming from crowded words. Psychol Sci 23(6):608–616

Tai-Hsiang Huang received his M.S. degree in Communication Engineering from National Taiwan University in 2008. He is currently working toward the Ph.D. degree in the Graduate Institute of Communication Engineering, National Taiwan University. His research interests are in the area of perceptual based image and video processing.

Su-Ling Yeh received the Ph.D. degree in Psychology from University of California at Berkeley, California, U.S.A. Since 1994, she has been with the Department of Psychology, National Taiwan University. She is the Recipient of Academic Award of the Ministry of Education and Distinguished Research Award of National Science Council of Taiwan, and is a distinguished professor of National Taiwan University. Her research interests lie in the broad area of multisensory integration and effects of attention on perceptual processes. Dr. Yeh serves as an advisory Council of the International Association for the Study of Attention and Performance (A&P). She also serves in the editorial board of Frontiers in Perception Science.

Yung-Hao Yang received the Master degree in General Psychology from National Taiwan University, Taipei, Taiwan, in 2008, where he is currently working toward the Ph.D. degree at the Graduate Institute of General Psychology. His research interests are in the area of visual perception, cross-modal Integration, and the neural mechanism of Consciousness.

Hsin-I Liao received the Ph.D. degree in psychology from National Taiwan University in 2009. She is currently working as a research associate at NTT Communication Science Laboratories. Her research interests lie in the broad area of preference decision making, effects of attention on perceptual processes, and neural mechanism of visual awareness.

Ya-Yeh Tsai received the M.S. degree in psychology from National Taiwan University in 2013. His research interests are in the area of multisensory integration, tactile perception and time perception.

Pai-Ju Chang received her B.S. degree in Psychology from National Taiwan University in 2013. She is currently working as a research engineer in Department of Computing and Information Science, Masdar Institute of Science and Technology. Her research interests focus on visual perception and social computing.

Homer H. Chen (S'83–M'86–SM'01–F'03) received the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign.

Since August 2003, he has been with the College of Electrical Engineering and Computer Science, National Taiwan University, Taipei, Taiwan, where he is Irving T. Ho Chair Professor. Prior to that, he held various R&D management and engineering positions with U.S. companies over a period of 17 years, including AT&T Bell Labs, Rockwell Science Center, iVast, and Digital Island (acquired by Cable & Wireless). He was a U.S. delegate for ISO and ITU standards committees and contributed to the development of many new interactive multimedia technologies that are now part of the MPEG-4 and JPEG-2000 standards. His professional interests lie in the broad area of multimedia signal processing and communications.

Dr. Chen was an Associate Editor of the IEEE Transactions on Circuits and Systems for Video Technology from 2004 to 2010, the IEEE Transactions on Image Processing from 1992 to 1994, and Pattern Recognition from 1989 to 1999. He served as a Guest Editor for IEEE Transactions on Circuits and Systems for Video Technology in 1999 and IEEE Transactions on Multimedia in 2011.