The Role of Dynamic and Static Visual Information in Real-World Visual Search

Authors: Pan Jing, Zhang Huiyuan, Chen Donghao, Xu Hongge, Pan Jing

Date: 2020-04-24T00:00:00+00:00

Abstract

Visual search in real-world environments is an essential ability for human and animal survival. Current visual search research predominantly employs static observers and stationary two-dimensional search targets, focusing on investigating the role of attention in search; existing theoretical models of visual search primarily summarize top-down attentional factors that affect search, while simplistically reducing bottom-up influences to image saliency. However, in real-world environments, observers or search targets can be in motion, and the visual information available during search includes dynamic optic flow and static image structure information. Previous research on visual recognition has found that combining these two types of information enables observers to accurately and persistently recognize scenes, events, and three-dimensional structures. By introducing these two types of visual information into existing theoretical models of visual search, we can reproduce search tasks in real-world environments and design experiments to investigate the visual search process that utilizes dynamic and static visual information, thereby improving existing visual search models. We believe that fully utilizing environmental information can improve search efficiency and has important application value in areas such as visual search training and intelligent search design.

Full Text

Preamble

The Role of Dynamic and Static Visual Information in Real-World Visual Search

PAN Jing; ZHANG Huiyuan; CHEN Donghao; XU Hongge
(Department of Psychology, Sun Yat-sen University, 132 Waihuan East Road, Guangzhou Higher Education Mega Center, Guangzhou, 510006, China)

Abstract

Visual search in real-world environments is a critical ability for human and animal survival. Current visual search research predominantly employs static observers and stationary two-dimensional search arrays, focusing primarily on the role of attention in search. Existing theoretical models of visual search have mainly summarized top-down attentional factors influencing search, while simplistically reducing bottom-up influences to image saliency. However, in real-world environments, both observers and search objects can move, and the visual information available during search includes dynamic optic flow and static image structure information. Previous research on visual recognition has demonstrated that combining these two types of information enables observers to accurately and persistently recognize scenes, events, and three-dimensional structures. This study introduces these two forms of visual information into existing visual search models to recreate real-world search tasks, designs experiments to investigate the visual search process utilizing dynamic and static visual information, and thereby refine current visual search models. We propose that fully exploiting environmental information can improve search efficiency and has important applied value for visual search training and intelligent search design.

Keywords: visual search; optic flow; biological motion; ecological perception theory

1.1 Research Background

Visual search refers to the process of finding a specific target among numerous items through observation and utilization of visual information. This is one of the key abilities for human and animal survival. Research on the visual search process and its cognitive mechanisms helps identify factors affecting search efficiency, thereby enabling targeted development of visual search training programs to improve search accuracy and speed. In classic visual search experiments, a search target and several distractors are randomly and independently presented on a computer screen (Treisman & Gelade, 1980; Koch & Ullman, 1987; Duncan & Humphreys, 1989; Wolfe & Gancarz, 1997), with targets and distractors distinguished by one or several dimensions—for example, finding a $ symbol among multiple letter S' s, or locating a red circle among geometric shapes of different colors. Compared to real-world visual search tasks, the search targets, distractors, and background information in such laboratory tasks are overly simplistic. In recent years, researchers have increased the ecological validity of visual search experimental paradigms by using static pictures of real objects as search arrays (either presented against a uniform color background or embedded in static scenes). These studies have found that factors such as scene structure/gist (Torralba, Oliva, Castelhano, & Henderson, 2006; Henderson & Hayes, 2017), search history (Ort, Fahrenfort, & Olivers, 2017; Wolfe, Cain, & Aizenman, 2019), target frequency (Wolfe & Wert, 2010; Wolfe, Boettcher, Josephs, Cunningham, & Drew, 2015), and target value (Clayton, Leonardo, Jan, & Geng, 2014; Ehinger & Wolfe, 2016) all influence visual search.

However, real-world visual search is far more complex than laboratory tasks. Specifically, in traditional visual search experimental paradigms, the observer, search objects, and search environment are mostly static, whereas in reality they can all move. Second, laboratory tasks typically use simple abstract two-dimensional images as targets and distractors, whereas real-world search objects are three-dimensional, far more complex, and may have countless feature combinations. More importantly, the appearance of objects in real scenes changes with viewing distance, viewing angle, lighting, movement of objects or observers, and can even be occluded (Foulsham & Underwood, 2009). Even when recent studies have used real photographs or virtual reality scenes as search environments, researchers still reduce rich three-dimensional dynamic and static information to two-dimensional projections of objects, using simple images as the starting point for the search process, resulting in at least half of the visual information (dynamic information) being excluded from existing visual search models.

Consequently, many researchers have questioned the validity of laboratory tasks, arguing that conclusions based on them cannot fully explain real-world behavior (Broadbent, 1991; Kingstone, Smilek & Eastwood, 2008) and may even be counterproductive, leading research in the wrong direction (Kingstone, Smilek, Ristic, Friesen, & Eastwood, 2003). Kingstone and colleagues (Kingstone et al., 2003) explicitly proposed that visual search research must explore new directions, integrating observers, objects, search tasks, and environments, particularly considering the observer' s state and natural environmental characteristics, and studying visual search from an embodied perspective. This aligns perfectly with the core ideas of Gibson' s ecological perception theory (Gibson, 1958; Gibson, 1979/1986).

The visual search process involves cognitive activities including perception (Treisman, 1982; Theeuwes, Kramer, & Belopolsky, 2004), attention (Árni, Ómar, & Thornton, 2014; Wolfe & Horowitz, 2017), working memory (Drew, Boettcher, & Wolfe, 2016; Drew, Boettcher, & Wolfe, 2017), and long-term memory (Woodman & Chun, 2006; Le-Hoa & Wolfe, 2015). These cognitive activities are not independent in the visual search process but rather interwoven and permeating, forming a continuous process. However, in most visual search research, discussions of the search process and mechanisms have focused on attentional guidance, allocation, or capture. In contrast, although perception plays an important role throughout the search process—from early feature registration (Treisman, Sykes, & Gelade, 1977; Treisman & Gormican, 1988; Wolfe, & Gray, 2007), through mid-stage target-noise separation (Eriksen & Schultz, 1979), to later serial recognition stages (Treisman & Gormican, 1988; Wolfe, Cave, & Franzel, 1989; Wolfe, & Gray, 2007)—research on perceptual information processing during search is surprisingly scarce. Nakayama and Martini (2011) proposed that perception research, particularly object recognition, can help us understand visual search. They argued that object recognition and visual search are essentially both pattern recognition. Object recognition requires using features from many dimensions to classify an object, while visual search tasks use features from a few dimensions to distinguish multiple objects. The two tasks are essentially the same, representing merely a trade-off between the number of dimensions and the number of objects. This shows that recognition and search are two ends of a continuum, closely connected, and therefore information that aids object recognition can also aid visual search. Based on this, the present study adopts an ecological perception perspective, proposes experimental concepts, investigates the impact of dynamic and static visual information on search, and supplements and improves existing visual search models.

1.2 Research Significance

At the theoretical level, this study breaks conventional thinking patterns, restores search to its most authentic context, and proposes that the states of observers and objects affect search behavior. This influence is information-based: different states of observers and objects produce different visual information, and different visual information leads to different search behaviors. Like other perceptual tasks, visual search is an active, dynamic, embodied process involving search objects, environmental backgrounds, observers themselves, and many other aspects. This study aims to improve the theoretical system of visual search, identify how stationary or moving observers in the real world use optic flow and image structure information to search for objects or events, and improve observers' ability to extract, integrate, and utilize visual information through training, thereby enhancing search efficiency.

In terms of application, humans live in three-dimensional environments that are constantly changing, and humans themselves are also in motion. How observers find targets while moving, and how they locate a moving target from a group of moving people, are important cognitive tasks with broad applications in real, dynamic environments. For example, in public safety, police need to quickly find targets in video surveillance, and emergency personnel need to rapidly locate people needing help in crowded, flowing areas—all requiring the use of dynamic and static visual information for search. Existing search theories are primarily based on static image information, with a process that roughly involves first "remembering" a target image, then finding the target among many images. However, when people move in an environment, the situation becomes complex; faces, shapes, and other image-based recognition cues are frequently occluded and easily altered. Visual search based solely on images is not the most effective approach and is prone to misses or false alarms. If motion information can be obtained and utilized during search, efficiency may be higher because motion patterns depend on physical characteristics (such as mass, length, friction coefficient, etc.; each person' s body shape is different, and walking posture is also different), and motion patterns are difficult to change. Therefore, both image structure information and optic flow information generated by motion are important in visual search. In this study, we verify whether observers can use visually generated motion information for search and improve cognitive models of search based on dynamic and static visual information.

The application value of this study' s results can be summarized as: (1) Developing targeted training programs that focus on training the extraction and integration of effective visual information to improve search accuracy and reduce search time; (2) Combining human behavioral research results with machine learning to make automated search efficient and low-cost. Existing intelligent search algorithms are mainly based on image information; in principle, the higher the image resolution and the more frames, the more accurate the search. However, such search entails heavy computational load and high requirements for processors, cooling systems, and batteries, making it unsuitable for small portable devices. In contrast, processing optic flow only requires low-frequency spatial signals and does not have high clarity requirements. Therefore, compared to traditional algorithms, optic flow-based search has advantages such as fast computation speed, low computational load, less heat generation, and low energy consumption, making it very suitable for portable search devices or drone search. This study breaks conventional thinking by introducing dynamic optic flow information into visual search, advancing applications in improving human search efficiency and artificial intelligence search algorithms, and holding high applied value and social significance. As Wolfe stated, "our health and safety rely, in part, on successful search" (Wolfe, 2003, p75).

2.1 Visual Search Theory and Development

Anne Treisman launched contemporary visual search research with her groundbreaking Feature Integration Theory (FIT). FIT divides the search process into two stages: a pre-attentive stage where features are automatically and processed in parallel across the visual field, followed by an attentive stage where attention binds features to enable serial recognition of objects (Treisman & Gormican, 1988). Duncan and Humphreys (1989) disagreed with Treisman' s dichotomy between parallel and serial search, instead proposing similarity theory. They argued that visual search tasks are easy when distractors are homogeneous and very different from the target, and difficult otherwise. To fill the gap in the attentional guidance mechanism of features in the pre-attentive stage, Wolfe, Cave, and Franzel (1989) modified FIT and proposed the Guided Search model. In this model, attentional guidance is divided into top-down and bottom-up components. Bottom-up guidance focuses on local contrast or physical saliency information of stimuli, while top-down guidance focuses on the degree of match between current items and the target across different features. Final search efficiency is a weighted sum of both types of guidance. The Guided Search model has been enormously influential, and almost all subsequent theoretical research on visual search has been refined and argued within this framework.

In recent years, with technological developments (particularly portable eye-tracking and virtual reality) and algorithmic advances (Bayesian estimation, network models, etc.), increasing numbers of visual search studies have enhanced the ecological validity of laboratory tasks by increasing the complexity of backgrounds and search objects (e.g., using pictures of real objects or scene photographs). Wolfe' s Guided Search model has also been continuously expanded (Wolfe herself has continuously improved the Guided Search model: Guided Search 2.0—Wolfe, 1994; Guided Search 3.0—Wolfe & Gancarz, 1997; Guided Search 4.0—Wolfe & Gray, 2007; Guided Search 5.0—Wolfe et al., 2015). In current models (see Figure 2 [Figure 2: see original paper]), top-down influences consist mainly of three components: template guidance, episodic guidance, and semantic guidance. Template guidance refers to the searcher' s knowledge and background information about the search target (Bahle, Matsukura, & Hollingworth, 2018; Duncan & Humphreys, 1989). Episodic guidance refers to where the target has previously appeared in similar contexts (Brooks, Rasmussen, & Hollingworth, 2010; Vo & Wolfe, 2012). Semantic guidance refers to where the target is likely to appear in similar contexts, influenced by background information, object-scene relationships, and object-object relationships (Wolfe, Cain, & Aizenman, 2019; see Wu, Wick, & Poumplun, 2014 for a review). To date, the analysis of top-down factors affecting search in real environments has been relatively well-developed, forming some consensus.

In the Guided Search model, bottom-up influences mainly include the image saliency of search objects (Koehler, Guo, Zhang, & Eckstein, 2014)—that is, the greater the image differences between search objects or the more prominent the target' s image, the faster and more accurate the search. Koch and Ullman (1987) proposed the saliency map, which forms a distribution map based on significant differences in multiple features of search objects to predict observers' search locations. This theory is supported by many experimental results (De Vries, Hooge, Wertheim, & Verstraten, 2013; Kamkar, Moghaddam, & Lashgari, 2018). However, recent studies have found that the influence of image saliency on attention allocation is limited to laboratory tasks and cannot generalize to real-world visual search. For example, when searching in real scenes, the saliency of visual stimuli cannot predict or explain observers' eye movements (Wu, Wick, & Pomplun, 2014). Given that eye fixation points are important behavioral indicators of attention allocation (Henderson & Hayes, 2017), this means that saliency map models cannot explain attention allocation in real search or predict search behavior in real environments (Foulsham & Underwood, 2009; Henderson, Malcolm, & Schandl, 2009).

We argue that saliency validated through laboratory tasks cannot explain real-world search because the settings for observers and objects in laboratory tasks are overly simplistic and idealized. First, some researchers have proposed that salient objects must enter the observer' s central visual field for their saliency to be effective (Wolfe, 2003; Foulsham, Chapman, Nasiopoulos & Kingstone, 2014). In search tasks conducted on laboratory computers, observers typically sit upright in front of the computer with their head fixed, and static search objects fall within the central visual field—clearly not the case in real-world search. For example, Foulsham et al. (2014) designed a real search task requiring participants to walk from the laboratory through several corridors into a mailroom containing a wall of identical-sized mailboxes (120 total), where they needed to find a target mailbox. In half the trials, researchers painted the target mailbox fluorescent pink, hoping the obvious color feature would make the target salient and trigger pre-attentive search. However, results showed that search reaction time was the same regardless of whether the target mailbox was painted pink; target saliency had no effect on reaction time. Researchers argued that in real scenes, observers are relatively small compared to the environment, so they must first move their body and head to scan the environment before using their eyes to search. Eye search is a secondary search nested within body search. When the head happens to be oriented toward the target direction and eye search begins, saliency can then play a role. Since the first stage of body search takes far longer than eye search (in the above experiment, body search took 26 seconds while eye search took 4 seconds), the benefit of saliency for real-world search cannot be manifested. Second, saliency is not static but dynamic and situated, changing with time and space. Existing theories hold that differences in color, shape, size, motion, and other dimensions of search objects affect search efficiency—the greater the differences, the more salient the target. However, in real environments, viewing time, observer posture, or object movement may change lighting, viewing angle, viewing distance, and occlusion relationships between objects, thereby altering the color, shape, size, and other image information of search objects. In such cases, the saliency of search objects is almost impossible to define or quantify. Additionally, real search may involve simultaneous changes in multiple dimensions—for instance, one object' s color may be prominent while another object is moving, making it difficult to determine which is more salient.

2.2 Ecological Perception Theory and Visual Information

In the visual search field, the typical approach to recreating real environments is searching within scene photographs. However, Gibson explicitly stated that real scenes and scene pictures are different: "Looking at Niagara Falls is not the same as looking at a photograph of Niagara Falls (Gibson, 1979)" ; similarly, searching in a picture of a kitchen is not the same as actually searching in a kitchen. The difference lies in the fact that in real environments, observers or objects can move. Motion continuously deforms objects' images and causes mutual occlusion, making search based purely on image matching unfeasible. However, motion generates another type of information: optic flow, which is dynamic information that can specify the structure and relationships of objects in the environment.

Ecological optics theory holds that all visual tasks depend on optical information. Gibson (1966) proposed that light enters the environment, is reflected by surfaces or objects, and forms ambient light. Ambient light carries information about the entire environment. For example, tiles, marble, and metal surfaces reflect different ambient light, so by detecting ambient light, people can know which is a kitchen wall, which is a countertop, and which is a sink.

Ambient light converges at a point of observation, forming an optic array. For a given observation point, the various surfaces constituting the static optic array have different visual solid angles, which correspond one-to-one with the layout structure of object surfaces in the environment, forming static image structure information. Static image structure information includes edges, shading, and contrast of color or intensity. This information is persistent—as long as objects exist, image structure information exists.

When the observer moves or objects in the environment move, the visual solid angles in the optic array also change: they may be added, disappear, enlarge, or shrink. The continuous change of the optic array forms optic flow information. The state of optic flow corresponds one-to-one with the observer' s relative motion speed, motion direction, and distance to moving objects in the environment —for example, objects farther from the observer have slower optic flow speed, and objects directly in front of the observer have faster optic flow speed than objects at the edge of the visual field. Optic flow is generated by motion and corresponds one-to-one with motion patterns; by detecting the state, direction, speed, and location of stationary points of optic flow, observers perceive their own or environmental objects' motion patterns. See Figure 1 [Figure 1: see original paper].

Figure 1: Summary of ecological optics theory. Gibson argued that observers use visual information in ambient light to accomplish perceptual tasks. Ambient light includes static image structure information and dynamic optic flow information.

Each object surface in the environment corresponds to a given observation point and forms a unique optic array, while the motion pattern of the observation point or environmental objects forms a unique optic flow. Such one-to-one correspondence is determined by natural laws. The image structure information projected from a surface in the environment to an observation point is constrained by geometric laws and is not random; the continuous optic flow generated by motion is constrained by dynamic and kinematic laws and is also not random. This regularity enables observers to accurately perceive the structure and properties of the environment through static and dynamic information.

Therefore, to recreate real-world search in the laboratory, using real pictures or virtual reality displays is insufficient; the key is to provide image structure and optic flow, which can build more ecologically valid experimental scenarios, set search conditions closer to real environments, and investigate visual search in real-world contexts.

2.3 The Role of Dynamic and Static Visual Information in Perceiving Scenes, Object Structure, and Events

In natural viewing environments, both dynamic and static visual information exist simultaneously. The combination of optic flow and image structure information can help accurately and stably perceive scenes, object structure, and events. First, research has found that observers can use optic flow generated by their own motion to recognize blurred scenes, with a positive correlation between optic flow intensity and scene recognition performance (Wu, Wang & Pan, 2019).

Second, observers can use dynamic visual information to accurately perceive the three-dimensional structure of objects through a mechanism called structure-from-motion (Domini, Vuong, & Caudek, 2002; Todd, Tittle, & Norman, 1995). For example, Lind and colleagues (Lee, Lind, Bingham, & Bingham, 2012) found that when cylinders with different width-to-depth ratios were placed before observers viewing from a 45° downward angle, observers could not perceive the three-dimensional structure of objects based solely on image structure when both the target and observer were stationary. However, as long as there was continuous perspective change of more than 45° between observer and target (either the observer or the target rotated more than 45°), observers could accurately perceive the three-dimensional structure of objects.

Third, optic flow and image structure can aid event recognition (Pan & Bingham, 2013; Pan et al., 2017). Events refer to objects in motion, and biological motion is one type of event that has been most intensively studied. Research has found that observers can recognize various movements, actor characteristics, and other non-biological motion events (e.g., a rolling ball, undulating water surface, etc.) through visual information generated by point-light motion (Bingham, Rosenblum, & Schmidt, 1995). Additionally, a few studies have added simple image information to biological motion paradigms (e.g., connecting points with lines or adding contour lines) and used Bayesian models ("ideal observer models") to analyze the information content in experimental stimuli, finding that changing the amount of visual information in stimuli affects the efficiency of recognizing and discriminating biological motion (Gold, Tadin, Cook, & Blake, 2008; Lu, Tjan, & Liu, 2017).

Scholars have studied the roles of dynamic and static visual information in event recognition and proposed the "kinematics-specified-by-dynamics theory" (Runeson & Frykholm, 1983). Researchers argue that because the physical dynamics behind each movement differ (e.g., the forces forming running, jumping, and walking are completely different), and each mover' s body has different physical properties (mass, limb length, joint flexibility, muscle strength, etc.), different movements by different movers have unique, fixed kinematic characteristics. When observers receive only dynamic visual information, they can perceive biological motion and the properties of the mover themselves based on observed motion features, thereby perceiving specific events. Bingham and colleagues subsequently proposed that the visual information specifying event dynamics is trajectory form—the relationship between a moving object' s position and velocity ( ), (trajectory form; Bingham, 1995; Bingham, Schmidt, & Rosenblum, 1995; Muchisky & Bingham, 2002; Wickelgren & Bingham, 2004, 2008b).

Trajectory form is influenced by forces. Each type of dynamics produces a unique trajectory form, so it can be used to specify events. Research has found that people are very sensitive to this dynamic information and can use it to distinguish very similar events, such as a hand-controlled pendulum versus a freely swinging pendulum (Muchisky & Bingham, 2002). More importantly, trajectory form information is not affected by viewing perspective; events can be recognized even from unfamiliar viewing angles (Wickelgren & Bingham, 2004, 2008b).

In summary, observer motion can generate optic flow; through motion, the structure of objects can be recovered to recognize stationary three-dimensional objects; by attending to objects' motion states, observers can recognize events and the properties of objects involved. Can the role of dynamic and static visual information in perceptual activities be transferred to visual search tasks? The answer is affirmative. As previously discussed, the biggest difference between real-world search and photograph-based search is motion. In real environments, relative motion between observer and object changes static image information and image saliency, so effective search mechanisms must be able to adapt to or resist changes in search object appearance caused by motion (Seidl-Rathkopf, Turk-Browne, & Kastner, 2015). Therefore, dynamic visual information independent of images and resistant to perspective changes (such as trajectory form) is likely the information people need in real search.

2.4 The Role of Dynamic and Static Visual Information in Visual Search Tasks

Biological motion paradigms containing only motion information have been applied in attention research (Ding, Yin, Shui, Zhou, & Shen, 2017; Myer, Vuong, & Thornton, 2015), but few studies have used biological motion paradigms to investigate visual search. These studies have found that observers can find a point-light walker among randomly moving points (Hirai & Hiraki, 2006), find an inverted walker among upright walkers (Wang, Zhang, He, & Jiang, 2010), find people walking in different directions (Cavanagh, Labianca, & Thornton, 2001), and distinguish different movements (Van Boxtel & Lu, 2011).

Real-world visual search typically involves both dynamic and static visual information simultaneously. However, early experimental tasks within the Guided Search model framework mostly used two-dimensional static graphics or symbols as search objects, and observers generally could not freely move their bodies to search. In real-world visual search, objects can be stationary or moving three-dimensional objects, and observers can also be stationary or moving. For example, in the aforementioned study by Foulsham et al. (2014), although participants walked into a mailroom to find a target mailbox, the experiment only manipulated bottom-up information by changing the target mailbox' s saliency (adding a pink border), without considering the interaction between dynamic and static information during movement. In recent years, some researchers have used different methods to investigate the impact of participant movement on search (Smith, Hood, & Gilchrist, 2008, 2010). For example, Ruddle and Lessels (2006) used simulation technology to design a task where participants searched for targets in a virtual reality scene, requiring them to find 8 targets at 16 different locations. The first group of participants could only sit in front of a screen (body fixed) and simulate turning and forward movement in the scene by moving a mouse; the second group could turn their bodies at a fixed position and use a stereo HMD to achieve turning in the scene, but still needed to move the mouse to advance to different locations; unlike the first two groups, the third group could walk to any location in the real environment to search. Results showed that participants with fixed bodies had poorer search efficiency than the other two groups, while participants allowed to walk freely had the highest search efficiency. This suggests that dynamic information generated by body movement combined with static information from search objects may facilitate improved visual search efficiency.

The "embodied memory model"proposed by Pan, Bingham, and Bingham (2013) suggests that when dynamic and static visual information simultaneously specify an event, optic flow has spatial accuracy and can calibrate image structure, helping observers accurately recognize the three-dimensional relationships between objects and the environment; image structure has temporal stability and can form embodied memory after motion stops and optic flow disappears, enabling observers to continuously perceive three-dimensional structure. We found that the combination of image information and optic flow information allows observers to accurately locate hidden or camouflaged targets. In studies by Pan et al. (2013, 2017), multiple targets were gradually occluded by distractors, and participants could use both types of visual information to accurately find targets during and after occlusion. In Pan, Bingham, Chen, and Bingham (2017), when targets and distractors had identical appearances but different spatial positions, participants could use both types of visual information to accurately and stably locate targets. According to Nakayama and Martini' s (2011) definition of recognition and search tasks (recognition involves identifying a target through multiple features, while search involves finding multiple targets through a few features), these two tasks actually lean more toward visual search.

In summary, existing research has shown that observers can use dynamic visual information to search for events (biological and non-biological motion) and static visual information to search for objects. Therefore, based on the existing Guided Search model, we add visual information variables and propose that optic flow and image structure are important bottom-up factors influencing search. Observers use differences in images between search objects to distinguish different individuals; objects with greater differences are more salient and easier to search for. Observers can use optic flow information to perceive search objects' motion features and underlying mechanical properties, thereby distinguishing different search objects (Figure 2). Of course, observers' own motion can generate not only optic flow information at the visual level but also proprioceptive and kinesthetic information. When observers conduct visual search while moving, motion provides more bottom-up visual information that aids visual search; however, we cannot exclude the influence of motion-related information on higher cognitive processes such as working memory and attention, which may be inhibitory (Mayer, Riddell, & Lappe, 2019).

Figure 2: Theoretical model proposed in this project. We argue that visual information has a bottom-up influence on real-world visual search, particularly optic flow and the interaction between optic flow and image structure. This study supplements the bottom-up role of optic flow information in visual search, but visually, kinesthetic, and proprioceptive information generated by motion may have top-down influences on visual search in other ways, which awaits future research. Black text represents the existing theoretical model, primarily based on Wolfe' s Guided Search model. Red text represents key concepts proposed in this theoretical framework.

3.1 Scientific Questions

The first scientific question this study aims to address is: What are the bottom-up factors influencing the search for three-dimensional objects and events in real environments? Existing theories divide factors influencing real-world search behavior into bottom-up and top-down categories, with bottom-up influences attributed to the image saliency of search objects. We argue this is insufficient. We propose that bottom-up influences on search are visual information, including static image structure and dynamic optic flow information, and that there is an interaction between the two types of information. Through Study 1 and Study 2, we investigate the roles of image structure information and optic flow information in searching for stationary three-dimensional objects and moving events, respectively, to answer the scientific question of how dynamic and static visual information are integrated and utilized to accomplish visual search in real environments.

The second scientific question this study aims to address is: Can traditional visual search theories generalize to and predict real-world visual search behavior? Traditional visual search research mostly uses two-dimensional images as search objects, and after decades of exploration has produced many theories that can explain two-dimensional image-based search behavior, such as radiologists identifying abnormal tissue from X-rays. However, real-world search tasks are more complex: search targets and distractors are three-dimensional, observers and objects can move, search perspectives change, background environments are complex, etc. Do visual search on flat images and visual search in real environments share similar behavioral patterns? We answer this question through two methods: directly comparing search performance (Study 1) and comparing the effects of search training (if training on flat images can improve real-world search performance, the two types of search are essentially similar; Study 3).

3.2 Research Plan

In this project, we investigate the process and mechanisms of visual search using dynamic and static visual information, as well as intervention methods to improve visual search efficiency, through three sub-studies (overall technical route see Figure 3 [Figure 3: see original paper]). The research primarily uses psychophysical methods and Bayesian estimation to analyze the amount of information in stimuli, search efficiency, and the relationship between them. Through three sub-studies, we systematically examine the roles of image structure and optic flow information in search, verify the bottom-up influence of visual information on search, and improve the search model based on visual information and attentional guidance. We then apply the theory to human visual search training and intelligent search design to help humans and machines better accomplish search tasks.

Figure 3: Overall technical route of this project. The project consists of three studies corresponding to static three-dimensional object search, event search, and visual search training. Green text indicates the visual information used, blue text indicates search tasks, and yellow text indicates study numbers.

3.2.1 Study 1: Visual Search for Stationary Target Objects Using Dynamic and Static Visual Information

Study 1 will explore how stationary or moving observers conduct visual search to find target objects when the search target is stationary, through three experiments addressing three questions: (1) Can structure-from-motion aid search for three-dimensional objects; (2) Is visual search efficiency affected by perspective change; (3) When observers move, perspective changes continuously, causing continuous changes in retinal images—can the two types of visual information resolve the impact of these changes on search.

In Experiment 1.1, we present stimuli on a computer screen using orthographic projection (0° viewing angle) and perspective projection (45° downward viewing angle) to compare search performance and determine whether perspective change affects visual search. In Experiment 1.2, we arrange real objects on a desktop, with participants sitting at the table observing the search array from a 45° downward angle, finding target objects under conditions where search objects are stationary, passively rotated, or actively rotated by participants. By comparing Experiments 1.2 and 1.1, we can determine whether search in real environments is the same as search in computer-simulated real scenes, and whether perspective change and structure-from-motion can facilitate search. In Experiment 1.3, we build more realistic and complex search scenarios in a virtual reality environment, allowing observers to move freely. Using conditions similar to Experiments 1.1 and 1.2, we compare results across three experiments to verify the validity of laboratory research and explore the impact of combining image structure and optic flow information on object search in real environments.

3.2.2 Study 2: Visual Search for Moving Events Using Dynamic and Static Visual Information

The main question of Study 2 is: When search objects are moving people (moving people are events), how do observers use visual information to find a particular person? Many studies using biological motion paradigms have shown that human observers are highly sensitive to human motion and can recognize human actions, classify actions, or distinguish gender, body size, emotion, etc., using only motion information (without image information). Can we find a specific moving individual from a group of moving people using only motion information? Additionally, in traditional biological motion paradigms, a moving person is simplified to a set of coordinated point-lights and presented from a sagittal view using orthographic projection, and in most cases, point-light groups are independent without overlap or interpenetration (e.g., one or several independent point-light walkers moving left or right on a computer screen). However, in real environments, multi-person movement is not limited to the frontoparallel plane but includes depth motion, with much overlap and occlusion. Moreover, the observer' s line of sight and the plane containing search objects may not be perpendicular, such as when viewing from above or surveillance footage typically having a downward viewing angle. When depth motion and inter-observer occlusion exist, and when there is perspective change between observer and search objects, can observers rely on motion information to search for a moving target? Third, in real environments, search objects have not only motion information but also image information. Previous studies have found that the presentation mode of biological motion (point-lights, connected lines, contour lines, or silhouettes) greatly affects action recognition efficiency, and adding some image information (e.g., connecting points with lines before presentation) makes recognition more efficient (Lu, Tjan & Liu, 2017). Meanwhile, image structure can also affect similarity between targets and distractors and between distractors, as well as the saliency of search objects, thereby changing search task difficulty. For these two reasons, what impact does adding image information have on searching for moving targets? Finally, in real environments, there is often more than one type of motion—for example, people walking and cars driving on a street. According to the "kinematics-specified-by-dynamics" theory, the dynamics underlying human walking and car driving are fundamentally different, with completely different mechanical properties that can be easily distinguished from motion information. At the information level, adding another completely different motion should not affect search. However, motion is a salient cue, and moving distractors can instantly capture observers' attention, affecting search efficiency. Therefore, what interactive influence do target saliency and kinematic features—two bottom-up influences—have on event search?

We will answer these questions through four experiments. Experiment 2.1 combines visual search and biological motion paradigms, presenting independent moving point-light groups from a sagittal direction using orthographic projection on a screen, requiring participants to find target movers among several moving point-light groups. In Experiment 2.2, when search objects walk interpenetrating in space, point-lights will overlap, interpenetrate, and exhibit non-rigid motion. Experimental materials will be created from two downward viewing angles (0° and 45°) to simulate search perspectives of observers immersed in crowds and surveillance footage. Participants find target events through point-light group motion. In Experiment 2.3, we use virtual reality to build a scenario of finding a person in a busy place, endowing search objects with image structure information, manipulating image saliency (e.g., changing the color or uniformity of searched crowds' clothing), and comparing event search performance when observers are stationary versus moving. Based on Experiment 2.3, Experiment 2.4 includes static and dynamic distractors to investigate the interactive influence of distractors' image saliency and motion characteristics on event search in real environments. We add moving distractors (e.g., adding moving vehicles to a search task for a pedestrian on a street) to explore the impact of irrelevant motion information on search; then we change the saliency of static distractors (e.g., adding flashing roadside signs) to explore the impact of image structure information saliency on search.

3.2.3 Study 3: Visual Search Training Using Simulation

Studies 1 and 2 investigate from a theoretical perspective how dynamic and static visual information are integrated and utilized to accomplish searches for stationary objects or dynamic events in real environments, and summarize the effects of various factors on search efficiency. Based on this theoretical foundation, Study 3 aims to identify effective training methods for improving visual search efficiency. The goal of search training is to improve accuracy and efficiency in complex search tasks, including stationary or moving observers searching for target objects or events. Training consists of four phases: pre-test, training, post-test, and retention. Pre-test measures baseline performance of untrained observers. Post-test measures performance after training. Retention examines whether higher search efficiency can be sustained for a period after training. The training phase is the most important stage, and an important principle in designing training programs is how practicing simple tasks can improve performance in complex tasks. Training must be based on factors affecting search identified through theoretical research, using simple, controllable, and targeted training tasks to improve performance in complex search tasks. Meanwhile, training effects can further demonstrate whether factors identified in theoretical research truly affect search.

Study 3 identifies the most effective training methods through three experiments. Each experiment' s pre-test, post-test, and retention phases involve tasks of stationary or moving observers searching for stationary targets and stationary observers searching for moving targets (corresponding to Studies 1 and 2), while the training phase is divided into three experiments according to the complexity of training content. Experiment 3.1: The training phase uses a virtual reality-built simulation environment identical to the other three phases. Experiment 3.2: The training phase uses abstract three-dimensional search objects but still contains optic flow and image structure information, such as the Lego block search or point-light walker tasks used in Studies 1 and 2. Experiment 3.3: The training phase uses traditional flat visual search paradigms (such as finding red circles among colored shapes) with different visual information but requiring the same attentional allocation and control. Through Experiment 3.1, we can determine whether visual search in real environments can be improved through training. Through Experiment 3.2, we can determine whether training targeting visual information can improve visual search performance. Through Experiment 3.3, we can determine whether training targeting attention during search can improve search performance. From Experiment 3.1 to Experiment 3.3, perceptual information gradually decreases while the search task principle remains consistent, so by comprehensively comparing the effects of the three training methods, we can indirectly understand the relationship between visual information factors and attentional factors affecting search, thereby indirectly verifying our proposed search theory based on visual information and attentional mechanisms.

4 Theoretical Construction

Successful visual search is a necessary skill for human survival and reproduction. A large body of visual search research has been conducted based on the Guided Search model. This model posits that during search, attentional guidance can be divided into top-down and bottom-up components. Top-down influences include template guidance, episodic guidance, and semantic guidance, which have been relatively well-studied with some consensus formed. The bottom-up component is simply reduced to "image saliency." However, recent studies have found that image saliency' s influence on attention allocation is limited to laboratory tasks and cannot generalize to real-world visual search (Henderson & Hayes, 2017; Wu, Wick & Pomplun, 2014). The key question becomes: What are the bottom-up factors influencing visual search? Additionally, traditional visual search research mostly focuses on stationary observers and stationary search objects presented in flat formats. However, this represents only one type of visual search. In real environments, search objects can be stationary or moving three-dimensional objects, and observers can also be stationary or moving. Whether traditional laboratory research results can generalize to and predict real-world visual search remains to be answered.

To address these two questions, our team has designed studies to provide answers. We integrate ecological perception theory (Gibson, 1966, 1979), introduce dynamic and static visual information to refine bottom-up influences, and propose a theoretical model (see Figure 2). We hope to achieve the following advances in visual search research:

First, Study 1 includes two situations: stationary observers searching for stationary three-dimensional objects (utilizing image structure information) and moving observers searching for stationary three-dimensional objects (utilizing image structure and global optic flow). Based on Study 1' s behavioral data, we will explore the influence of dynamic and static visual information on searching for three-dimensional objects in real environments and understand the behavioral patterns of three-dimensional object search. Then, according to Pan, Bingham, and Bingham' s (2013, 2017) "embodied memory model" theory, we will compare flat structure-based search with real-environment three-dimensional object search to verify whether theories derived from traditional flat search research apply to three-dimensional object search. The results of this study can, on one hand, demonstrate the role of optic flow in visual search, filling gaps in previous theoretical models. On the other hand, they can also show that image structure can preserve objects or events specified by optic flow, giving search persistence.

Second, Study 2 will design two tasks: stationary observers searching for moving three-dimensional objects (i.e., events, utilizing image structure and local optic flow) and moving observers searching for events (utilizing image structure and global and local optic flow). Combining biological motion paradigms, we will verify the role of dynamic and static visual information in event search in real environments. We will verify whether optic flow' s specification of motion features and their underlying mechanical properties enables observers to distinguish events and achieve accurate, persistent visual search through combination with image information. Additionally, after adding different perturbations (e.g., perspective changes, occlusion), we can explore the transformation resistance of optic flow information in event specification. Finally, we will provide static image structure and optic flow information in virtual simulation environments to understand search processes and patterns in complex scenes.

Finally, based on the first two studies, Study 3 will identify effective training methods for improving visual search efficiency. The first experiment verifies whether repeated practice of visual search in virtual reality scenes can improve search efficiency; the second experiment emphasizes training the ability to extract and integrate image structure and optic flow information, testing whether practicing search for moving three-dimensional objects can improve performance in complex search tasks in virtual simulation scenes. The third experiment uses traditional visual search paradigms, testing whether practicing search for flat graphics or symbols can improve performance in complex search tasks in virtual simulation scenes. If training effects differ between Experiments 2 and 3, this indirectly indicates that traditional visual search paradigms cannot fully represent visual search in three-dimensional environments.

In summary, this study will use multiple research techniques and methods to systematically examine the roles of image structure and optic flow information in search, investigate the bottom-up influence of dynamic and static visual information on search, and improve the search model based on visual information and attentional guidance. Based on this study, future research can further explore the top-down influence of visual, proprioceptive, and kinesthetic information generated by observers' own motion on search tasks, thereby improving a visual search theoretical system with high ecological validity. For example, foraging tasks are considered a natural task similar to visual search. In foraging tasks, observers can move their bodies to repeatedly view and find targets (Ehinger & Wolfe, 2016; Wolfe, Cain, Ehinger, & Drew, 2015), where the interaction between motion-related information and higher cognitive functions (attention, memory, etc.) becomes particularly important. On the other hand, ecological perception theory does not distinguish between biological and non-biological motion; both are events that can be recognized by capturing dynamic trajectory form information. However, our currently planned visual search experiments only use biological motion as search objects. Future research can expand search objects to other types of events to verify the universality of the theory of searching for events through dynamic visual information.

Árni Kristjánsson, Ómar I. Jóhannesson, & Thornton, I. M. (2014). Common attentional constraints in visual foraging. Plos One, 9(6), e100752.

Bahle, B., Matsukura, M., & Hollingworth, A. (2018). Contrasting gist-based and template-based guidance during real-world visual search. Journal of Experimental Psychology: Human Perception and Performance, 44, 367–386.

Bingham, G. P. (1987). Kinematic form and scaling: Further investigations on the visual perception of lifted weight. Journal of Experimental Psychology: Human Perception and Performance, 13(2), 155.

Bingham, G. P., Schmidt, R. C., & Rosenblum, L. D. (1995). Dynamics and the orientation of kinematic forms in visual event recognition. Journal of Experimental Psychology: Human Perception and Performance, 21(6), 1473.

Broadbent, D. E. (1991). A word before leaving. In D. E. Meyer & S. Kornblum (Eds.), Attention and performance XIV (pp. 863–879). Cambridge, MA: Bradford Books/MIT Press.

Brooks, D. I., Rasmussen, I. P., & Hollingworth, A. (2010). The nesting of search contexts within natural scenes: Evidence from contextual cuing. Journal of Experimental Psychology: Human Perception and Performance, 36, 1406–1418.

Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: Sprites. Cognition, 80(1-2), 47-60.

Clayton, H. , Leonardo, C. , Jan, T. , & Geng, J. J. . (2014). Reward-priming of location in visual search. PLoS ONE, 9(7), e103372-.

De Vries, J. P., Hooge, I. T., Wertheim, A. H., & Verstraten, F. A. (2013). Background, an important factor in visual search.. Vision Research, 86, 128–138.

Ding, X., Yin, J., Shui, R., Zhou, J., & Shen, M. (2017). Backward-walking biological motion orients attention to moving away instead of moving toward. Psychonomic Bulletin & Review, 24(2), 447–452.

Domini, F., Vuong, Q. C., & Caudek, C. (2002). Temporal integration in structure from motion. Journal of Experimental Psychology: Human Perception and Performance, 28(4), 816.

Drew, T., Boettcher, S. E. P., & Wolfe, J. M. (2016). Searching while loaded: Visual working memory does not interfere with hybrid search efficiency but hybrid search uses working memory capacity. Psychonomic Bulletin & Review, 23(1), 1–12.

Drew, T., Boettcher, S. E. P., & Wolfe, J. M. (2017). One visual search, many memory searches: An eye- tracking investigation of hybrid search. J Vis, 17(11), 5.

Duncan, J. S., & Humphreys, G. W. (1989). Visual search and stimulus similarity. Psychological Review, 96(3), Ehinger, K. A. , & Wolfe, J. M. . (2016). When is it time to move to the next map? Optimal foraging in guided visual search. Attention, Perception, & Psychophysics, 78(7), 2135–2151.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. Perception & psychophysics, 25(4), 249–263.

Foulsham, T., Chapman, C. S., Nasiopoulos, E., & Kingstone, A. (2014). Top-down and bottom-up aspects of active search in a real-world environment. Canadian Journal of Experimental Psychology, 68(1), 8–19.

Foulsham, T., & Underwood, G. (2009). Does conspicuity enhance distraction? Saliency and eye landing position when searching for objects. The Quarterly Journal of Experimental Psychology (2006), 62, 1088–Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. British journal of psychology, 49(3), 182–194.

Gibson, J. J. (1986). The ecological approach to visual perception. Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc. (Original work published 1979).

Gold, J. M., Tadin, D., Cook, S. C., & Blake, R. (2008). The efficiency of biological motion perception. Attention Perception & Psychophysics, 70(1), 88–95.

Henderson, J. M., & Hayes, T. R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. Nature Human Behaviour, 1(10), 743–747.

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. Psychonomic Bulletin & Review, 16(5), 850–856.

Hirai, M., & Hiraki, K. (2006). Visual search for biological motion: An event-related potential study.

Neuroscience Letters, 403(3), 299–304. Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. Attention Perception & Psychophysics, 14(2), 201–211.

Kamkar, S., Moghaddam, H. A., & Lashgari, R. (2018). Early Visual Processing of Feature Saliency Tasks: A Review of Psychophysical Experiments. Frontiers in Systems Neuroscience, 12.

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive ethology: A new approach for studying human cognition. British Journal of Psychology (London, England: 1953), 99, 317–340.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention, researchers! It is time to take a look at the real world. Current Directions in Psychological Science, 12, 176 -180.

Koch, C., & Ullman, S. (1987). Shifts in selective visual attention: towards the underlying neural circuitry. Human neurobiology, 4(2), 115-141.

Koehler, K., Guo, F., Zhang, S., & Eckstein, M. P. (2014). What do saliency models predict. Journal of Vision, 14(3), 14-14.

Lee, Y.L., Lind, M., Bingham, N. & Bingham, G.P. (2012). Object recognition using metric shape. Vision Research, 69, 23-31.

Le-Hoa, V. M., & Wolfe, J. M. (2015). The role of memory for visual search in scenes. Annals of the New York Academy of Sciences, 1339(1), 72-81.

Lu, H., Tjan, B. S., & Liu, Z. (2017). Human efficiency in detecting and discriminating biological motion.

Journal of Vision, 17(6), 4-4. Mayer, K. M., Riddell, H., & Lappe, M. (2019). Concurrent processing of optic flow and biological motion. Journal of Experimental Psychology: General.

Mayer, K. M., Vuong, Q. C., & Thornton, I. M. (2015). Do People "Pop Out" ?. PLOS ONE, 10(10).

Muchisky, M.M. & Bingham, G.P. (2002). Trajectory forms as a source of information about events. Attention Perception & Psychophysics, 64(1), 15-31.

Meijer, F., & Van der Lubbe, R. H. (2011). Active exploration improves perceptual sensitivity for virtual 3D objects in visual recognition tasks. Vision Research, 51(23-24), 2431-2439.

Nakayama, K., & Martini, P. (2011). Situating visual search. Vision Research, 51(13), 1526-1537.

Ort, E. , Fahrenfort, J. J. , & Olivers, C. N. L. . (2017). Lack of free choice reveals the cost of having to search for more than one object. Psychological Science, 095679761770566.

Pan, J. S., Bingham, N., & Bingham, G. P. (2013). Embodied memory: Effective and stable perception by combining optic flow and image structure. Journal of Experimental Psychology: Human Perception and Performance, 39(6), 1638-1651.

Pan, J. S., Bingham, N., & Bingham, G. P. (2017). Embodied memory allows accurate and stable perception of hidden objects despite orientation change. Journal of Experimental Psychology: Human Perception and Performance, 43(7), 1343-1358.

Pan, J. S., Bingham, N., Chen, C., & Bingham, G. P. (2017). Breaking camouflage and detecting targets require optic flow and image structure information. Applied Optics, 56(22), 6410-6418.

Pan, J. S., Li, J., Chen, Z., Mangiaracina, E. A., Connell, C. S., Wu, H., ··· & Hassan, S. E. (2017). Motion- generated optical information allows event perception despite blurry vision in AMD and amblyopic patients. Journal of Vision, 17(12), 13-13.

Rieser, J. J., Pick, H. L., Ashmead, D. H., & Garing, A. E. (1995). Calibration of human locomotion and models of perceptual-motor organization. Journal of Experimental Psychology: Human Perception and Performance, 21(3), 480.

Ruddle, R. A., & Lessels, S. (2006). For efficient navigational search, humans require full physical movement, but not a rich visual scene. Psychological Science, 17(6), 460-465.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person- and-action perception: expectation, gender recognition, and deceptive intention. Journal of Experimental Psychology General, 112(4), 585-615.

Seidlrathkopf, K. N., Turkbrowne, N. B., & Kastner, S. (2015). Automatic guidance of attention during real- world visual search. Attention Perception & Psychophysics, 77(6), 1881-1895.

Smith, A. D., Hood, B. M., & Gilchrist, I. D. (2008). Visual search and foraging compared in a large-scale search task. Cognitive processing, 9(2), 121-126.

Smith, A. D., Hood, B. M., & Gilchrist, I. D. (2010). Probabilistic cuing in large-scale environmental search. Journal of Experimental Psychology: Learning, Memory, and Cognition, 36(3), 605.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision:

Reinterpreting salience. Journal of Vision, 11(5), 5-5.

Theeuwes, J. , Kramer, A. F. , & Belopolsky, A. V. . (2004). Attentional set interacts with perceptual load in visual search. Psychonomic Bulletin & Review, 11(4), 697-702.

Todd, J. T., Tittle, J. S., & Norman, J. F. (1995). Distortions of three-dimensional space in the perceptual analysis of motion and stereo. Perception, 24(1), 75-86.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. Psychological review, 113(4), Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. Journal of Experimental Psychology, 8(2), 194.

Treisman, A. and Gelade, G. (1980) A feature-integration theory of attention. Cognitive Psychology, 12, 97-Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries.

Psychological Review, 95(1), 15-48. Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. Attention and performance VI, 333.

Van Boxtel, J. J., & Lu, H. (2011). Visual search by action category. Journal of Vision, 11(7), 19-19.

Vo, M. L., & Wolfe, J. M. (2012). When does repeated search in scenes involve memory? Looking at versus looking for objects in scenes. Journal of Experimental Psychology: Human Perception and Performance, 38(1), 23-41.

Wang, L., Zhang, K., He, S., & Jiang, Y. (2010). Searching for life motion signals: Visual search asymmetry in local but not global biological-motion processing. Psychological Science, 21, 1083-1089.

Wickelgren, E.A. & Bingham, G.P. (2004). Perspective distortion of trajectory forms and perceptual constancy in visual event identification. Attention Perception & Psychophysics, 66, 629-641.

Wickelgren, E. & Bingham, G.P. (2008). Trajectory forms as information for visual event recognition:3D perspectives on path shape and speed profile. Attention Perception & Psychophysics, 70(2), 266-278.

Wolfe, J.M. (1994) Guided Search 2.0: A revised model of visual search. Psychonomic Bulletin &. Review, 1, Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. Trends in Cognitive Sciences, 7(2), 70-76.

Wolfe, J. M. (2012). Guided search 4.0: Current progress with a model of visual search. Integrated Models of Cognitive Systems, 1(2), 202-238.

Wolfe, J. M. , Boettcher, S. E. P. , Josephs, E. L. , Cunningham, C. A. , & Drew, T. . (2015). You look familiar, but i don' t care: Lure rejection in hybrid visual and memory search is not based on familiarity. Journal of Experimental Psychology: Human Perception and Performance, 41(6), 1576-1587.

Wolfe, J. M., Cain, M. S., Ehinger, K. A., & Drew, T. (2015). Guided Search 5.0: Meeting the challenge of hybrid search and multiple-target foraging. Journal of Vision, 15(12), 1106-1106 Wolfe, J. M., Cain, M. S., & Aizenman, A. M. (2019). Guidance and selection history in hybrid foraging visual search. Attention, Perception, & Psychophysics, 81(3), 637-653.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception and Performance, 15(3), Wolfe, J. M., & Gancarz, G. (1997). Guided Search 3.0. In Basic and clinical applications of vision science (pp. 189-192). Dordrecht: Springer.

Wolfe, J. M., & Gray, W. (2007). Guided search 4.0. Integrated models of cognitive systems, 99-119.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. Nature Human Behaviour, 1(3), 0058.

Wolfe, J. M., & Wert, M. J. V. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. Current Biology, 20(2), 121–124.

Woodman, G. F., & Chun, M. M. (2006). The role of working memory and long-term memory in visual search.

Visual Cognition, 14(4-8), 808-830. Wu, C., Wick, F. A., & Pomplun, M. (2014). Guidance of visual attention by semantic information in real-world scenes. Frontiers in Psychology, 54–54.

Wu, H., Wang, X. M., & Pan, J. S. (2019). Perceiving blurry scenes with translational optic flow, rotational optic flow or combined optic flow. Vision Research, 158 (2019) 49–57.

Visual Search in Real World: The role of dynamic and static optical information PAN Jing; ZHANG Huiyuan; CHEN Donghao; XU Hongge (Department of Psychology, Sun Yat-sen University, GuangZhou, 51006, China) Visual search is a ubiquitous task and a critical skill for men and animals. Existing studies on visual search mainly focus on attentional guidance and the top-down cognitive influences on search effectiveness. The bottom-up influence on visual search is, rather crudely, simplified as objects' image saliency. However, when searching in real world, where the observer and/or objects move, both static image information (the saliency of which has been considered in existing search models) and dynamic optic flow information are available.

Optic flow is generated by the relative motions between an observer and world objects. So by detecting flow patterns, observers get to know the kinematic properties of events (which is defined as objects in motion) and hence perceive the physical properties of constituent objects, such as the mass, size and frictional coefficient etc.. These physical properties distinguish objects and allow the observer to search for a particular one. We integrate dynamical perceptual information (i.e. optic flow) into existing search models and in two studies, we test how combined dynamical and static perceptional information affect visual search for three-dimensional objects and for moving people, when the observer is stationary or moving. Furthermore, we attempt to develop a training protocol that improves search effectiveness in real world. Findings from this project will bring forth new theories for understanding visual search in real world, and have direct applications on personnel training and intelligent search designs.

Keywords: visual search; optical flow; biological motion; ecological theory of perception

Note: Figure translations are in progress. See original paper for figures.