Visual Attention and Eye Movements

Kang-chen Chen, Hye Jung Choi
Donald Bren School of Information and Computer Sciences (ICS)
University of California, Irvine
kangchec@uci.edu hchoi7@uci.edu

1. Introduction

Visual perception occurs when the eyes retrieve information from the visual field. The process of choosing "what" to observe is selective. Our eyes do not have enough attention to focus on every object in the field and gather their information; therefore, we need to select what is important to process. There are many different characteristics to visual selection such as when it occurs and how it occurs.

Visual selection can either be overt or covert. Overt selections are generally performed by eye movement. Our eyes will move to locations of interest to retrieve information. Covert selections are performed by visual attention. Even when our eyes are fixated on one location, we can still shift our attention to observe the different properties of that same location (i.e. color and shape).

2. Eye Movements

There are two major functions to the eye, fixation and tracking. Fixation is to position the target object into the fovea. This allows our eyes to maximize the focus we can give to the object. Tracking is the ability to fixate on objects even when they are moving. This is important because most real world objects move, and without the ability to track, we will have a very difficult time perceiving anything.

One common mistake people make is that the eyes simultaneously process all information in our visual field, but this is not true. There is simply too much information to process. Visual perception occurs over time, and each location and objects in the field are observed in sequence with a combination of eye movements and shifts in visual attention.

2.1. Eye Movement Types

There are six different types of eye movement. Each type of movement serves a different function in visual perception.

2.1.1. Physiological Nystagmus

Physiological nystagmus are tiny, involuntary eye movements that occurs constantly. These movements are not detectable under normal circumstances; however, scientists have devised ways to create situations where the eye functions as stationary. For example, subjects can wear a special contact lens that has a small light attached to it. This will ensure that the light will constantly be pointed at the same part of the eye. Another method is to use a complex graphics system, which tracks the movements of the eye. As soon as the eye moves, the image would also move, which creates the illusion that the eye is not moving.

The results from these experiments are surprising. The most impressive discovery is that the stabilized image would disappear after some time. This indicates that the visual system primary works using moving edges and reconstructing the visual world from that information.

Physiological nystagmus is the only type of movement which does not have a selective function. However, it is a necessary part of the visual system because without it, our eyes will not be able to detect anything.

2.1.2. Saccadic Movements

Saccadic movements are very rapid, ballistic movements that brings new objects of interest into the fovea. Saccades are said to be ballistic because once it starts, it will not stop until it reaches the target location. Even if the target location moves or disappears, the eyes would immediately find a new target and saccade to that point.

Saccades are also very quick. A single saccade only takes about 150-200 ms to plan and execute. The act of movement is also very fast, reaching a max speed of 900 degrees per second. The quick movements allow the eyes to concentrate on various parts of the visual world to gather information and paint a grand picture in our brains.

Images at saccade fixations are clear; however, the images during saccades are normally missed. This phenomenon is called saccadic suppression. Experiments have shown that our eyes actually do catch the images between during saccades; however, the images are blurry compared to the images perceived at fixations. The sharp images dominates over the blurry images, so we only consciously perceive the images at fixations.

2.1.3. Smooth Pursuit Movements

Smooth pursuit movements are used to track moving objects. This is an important function because most objects move. In the past, preys need to track moving predators to stay alive, and predators will need to track moving preys to catch them. Without the ability to track moving objects, we would not even be able to safely walk on the streets because we won't be able to perceive cars, bikes, etc.

There are four primary differences between smooth pursuit movements and saccades. First, smooth pursuit movements are smooth and continuous, but saccades move in a jerky and abrupt fashion. Second, smooth pursuit movements require constant feedback to track the moving object. Third, smooth pursuit movements are slower than saccades. Saccades move at 900 degrees per second, but smooth pursuit movements can only move at 100 degrees per second. Finally, the amount of clarity is different. The tracked object is clear, however, everything in the environment which does not move at the same direction or speed would appear smeared.

The speed of the moving object determines how well we can track the object. The faster it moves, the harder it is for our eyes to catch up. People are born with an innate ability to track object; however, through practice, the ability could be improved. For example, a study shows that baseball players generally have better dynamic acuity because they are trained to track baseballs as it is being pitched at a high speed.

2.1.4. Vergence Movements

Vergence movement's purpose is to select the distance of the object we want to observe. Saccades and smooth pursuit movements help us focus the object at the fovea, but without vergence movement, the eyes will not converge to the point of interest. Our eyes will converge less for farther objects and converge more for closer objects.

Vergence movements are very slow compared to other eye movements. The speed would almost never exceed 10 degrees per second. Another difference between vergence and other eye movements is that it is the only disconjugate eye movement. This means that during vergence movements, the eyes move in different direction. Saccades, smooth pursuit, and other movements that will be discussed shortly are all conjugate eye movements (moves in the same direction).

Vergence movement often works with saccades and smooth pursuit movements. During a saccade, if the new fixation is at a different depth from the previous fixation, then vergence movement would also occur to correct the angle between the eyes to focus on the right depth. This is also true for smooth pursuit movements. If the tracked object changes depth, then vergence movement would initiate to correct the difference in angle.

2.1.5. Vestibular Movements

The three previous selection eye movements are all used under one condition, when the head is still. Vestibular works together with optokinetic movements to retain the target object on the fovea when the head moves. This movement will move the eyes in the opposite direction of the head to keep the eyes focused on the same point at all times.

Vestibular movements are controlled by the vestibular system in the inner ear, which provides information on the changes in position and orientation of the head. When the head moves, the vestibular system would send signals to the eyes to compensate for the movements.

Although vestibular movements are slower than saccades, they are faster and more accurate than smooth pursuit movements. A simple way to test this is to stare at a stationary object and move your head left and right. You can start slowly and increase the speed, and you will find that the object will remain clear even if your head rotates very quickly. Next, keep your head still, and move the object. Start slow, and increase the speed. You will find that the object becomes blurry when it is moving at a moderate speed.

The ability to track objects while the head moves is more important than it appears. Similar to smooth pursuit movements, it is actually essential for our survival. Imagine that you are walking or running outside. Without vestibular movements, your eyes cannot focus while you move, making it extremely dangerous to travel. The amount of blur will be great enough to hinder recognition of signs and faces.

2.1.6. Optokinetic Movements

Optokinetic movements work together with vestibular movements. This type of movement only occurs when a large portion of the visual field moves across the retina. Besides the difference in functionality, optokinetic movements also have a different source of information. Vestibular movements are driven by the vestibular system, while optokinetic movements are guided by the transitions in the entire visual field.

2.2. The Physiology of the Oculomtor System

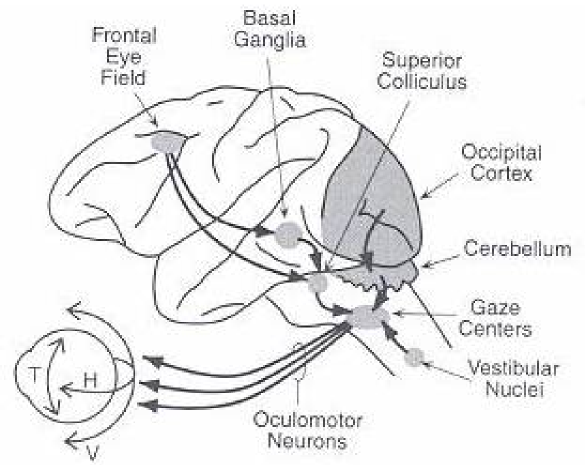

Figure 2.1 Brain areas that control eye movements

A common misconception that people have is that the eyes are controlled by a central location in the brain; however, that is not the case. Each type of eye movements is controlled by a different part of the brain, and different signals are sent to control the movements. There are still a lot of mystery surrounding brain mechanics, thus there may be hidden properties and connections between the brain and the eyes that have not been discovered yet (i.e. the precise mechanisms of each neural center)

Although the actual movements are initiated from different parts of the brain, there are only three pairs of muscles that perform all the movements. One set of muscles control the horizontal movements, one set controls the vertical movements, and the last set of muscles control the torsional movement (clockwise and counterclockwise). The gaze center controls these muscles; therefore, it is also known as the final common pathway of the oculomotor system.

Below is a list of eye movements, followed by the part of the brain that controls it:

- Physiological Nystagmus – Not controlled by the brain. They are tremors of the eye muscles.
- Saccades – controlled by the frontal eye fields in the frontal cortex
- Smooth pursuit movements – controlled by information from the motor channels in the visual cortext
- Vergence movements – controlled by visual feedback and occipital cortex
- Vestibular movements – driven by three-neural reflex arc that begins in the vestibular system
- Optokinetic movements – controlled by the critical motion pathway and subcortical pathway

3. Saccadic Exploration of the Visual Environment

In this section, we will discuss more about saccades because its functionalities are important for visual selection of complex images. It is obvious that focusing the eyes on one point could only yield so much information. It is imperative for our eyes to fixate on different locations to gather information and weave them into a bigger picture. Saccades allow us to quickly move our eyes to various locations, moving the objects of interest into the center of our fovea, and perceive with the highest possible resolution.

3.1. Patterns of Fixation

One important aspect of saccadic exploration is the pattern or sequence in which the eyes fixate. Studies have shown that the eyes will make fixations to

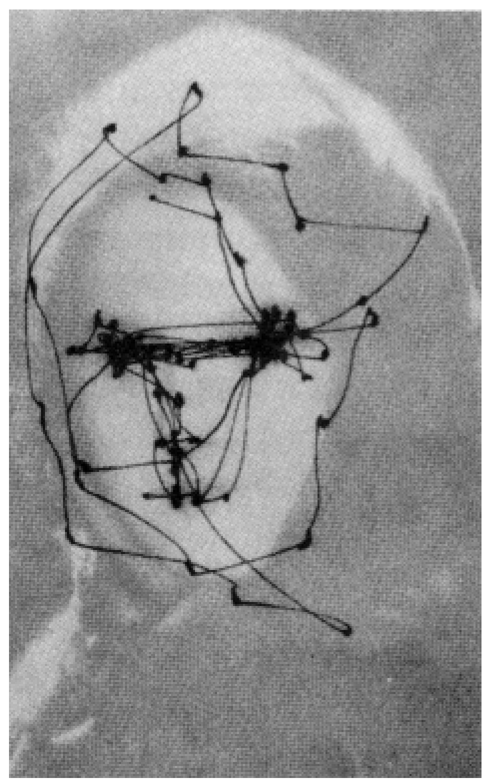

Figure 3.1 Image with sequences of fixation

points that are most informative. For example, in Figure 3.1, most of the fixations are made around the eyes, nose, and mouth. Those are distinct features that could help recognize who the person is. The edges of the face are also part of the sequence of fixation. The edges help define the shape of an object, which could then be used to figure out what the object is.

There are several factors that determine the pattern of fixation. One factor is the person's task or goals. For example, if a person was shown Figure 3.2 and asked if there is an airplane, the first places they would fixate on will be the sky and the clouds. However, if they were asked if there is anyone on the streets, they will fixate to the side walks and the roads first.
Modern tools have allowed scientists to track not only where the fixations occur, but also in what sequence. Recurring sequences of fixation in which a person saccades is called scan path.

Figure 3.2

3.2. Transsaccadic Integration

The information gathered from saccades are various separate images, but the results on the retina is actually a single, unified image. Therefore, there must some process that integrates these images and combine them into one entity. The entity could then be used by the brain to perceive the visual field. There are two different theories that attempt to explain this phenomenon.

The first theory is spatiotopic fusion hypothesis, which states that the contents of the sequences of saccades are mappend into a larger, spatially organized memory array and integrated based on its location. Although this hypothesis sounds logical, experiments have proven this to be inaccurate. Integration is actually performed at a more abstract level.

The second theory, schematic map, addresses the abstract integration problem. The schematic map is a representation of possible samplings of scenes that are integrated together by the expected results of seeing the scenes. The objects are encoded with distances and directions that will lead to other objects or features. This means that, given the same object and set of relations of objects, direction, and distance as input, the eyes would believe that a given object exists when it really doesn't.

4. Visual Attention

When our eyes fixate to different locations in a visual field, they do not process every single piece of information available. Instead, the eyes will go through a process to select different aspects to retrieve. Our eyes are capable of different levels of retrieval. It could either retrieve a large amount of low level information from the entire visual field, or it can retrieve high level details from the center of the fovea. We can also concentrate on a single object, saccade to various locations, and gather information mainly about that object.

The process could be split into two phases, recruiting and focusing. Recruiting is the act of gathering visual attention and other resources. Focusing is the act of using these resources to select aspects of visual information. Each of the two visual phases has a corresponding property. Recruiting is concerned with capacity, which is the amount of perceptual resources. The capacity is constantly changing. Some factors include alertness, motivation, and time of the day. Focusing is concerned with selectivity. Since there is a limited amount of capacity for attention, the eyes need to select which objects to process. Selectivity is studied more often than capacity; therefore, the remainder of this paper will be focused on selectivity.

Visual selection occurs in multiple phases. The first phase is a overt action, where our eyes use low level information to locate areas of interest. After the eyes have chosen and moved to a location of interest, an covert selection occurs. Although our eyes do not move, our attention actually shifts to various aspects within our field of vision. One situation where overt visual selection occurs is when people go out shopping. For example, if a person was presented with Figure 4.1. There attention could be focused on different aspects even if their eyes remain stationary. They could either concentrate on color, shape, or words.

There are two types of visual selection, spatial selection and property selection. Spatial selection focuses attention on a fixed region. Various aspects and objects within the filed will be examined. One way to test spatial selection is to turn off all the lights, and then turn them on for a very short amount of time.

During the limited amount of visibility, even though our eyes do not have time to select objects and process information in detail, it is enough to generate an image on the retinas with visual information. Property selection is the selection of features of an object. The example given in the previous paragraph is also an example of property selection.

Visual selection is given based on the importance of the objects in the visual field. Some objects are innately more important. For example, moving objects are essential for survival; therefore, they naturally have a higher priority than stationary objects. The level of importance of object is also based on personal experience. For example, a person's name is normally more important than other words they hear. It is common for a person to hear their name even if they are standing in the middle of a party.

Figure 4.1

4.1. Early versus Late Selection

So far we have discussed what is visual selection, and now we will focus the topic on when visual selection actually occurs. Since selection is based on the importance of objects in the visual field, it is logical to believe that visual selection occurs late. However, without first selecting an area to perceive, no information would be available. Likewise, if visual selection occurs early, the selection would be essentially random because there is no information available to make decision. This is called the paradox of intelligent selection. We will shortly describe experiments that were conducted to determine when selection actually occurs.

4.1.1. Auditory Attention

Researchers studied auditory attention to determine whether attention operates early or late. Although this research is not related to visuals, the results can provide indications of how human attention works. Subjects for this research were given a shadowing task, which they had to repeat the messages that come through one of the ears. The subjects were told to listen to one side, and then asked questions on the unattended side. The results showed that humans can perceive low level information such as whether the speaker was male or female, but was unable to even determine the language in which the message was spoken.

The next research came from British psychologist Donald Broadbent, who proposed that auditory attention operates early. He believes that the ears only analyze gross sensory features, and would select one ear for further process. This theory is called the filter theory. However, further research showed that this process was more complex than originally proposed.

Auditory attention actually occurs both early and late in the auditory process. The new theory that was proposed is the attenuator theory, which states that the unattended channel is not blocked, but attenuated. This means that the sounds coming through the unattended channel is reduced in importance. After this initial phase, the selection operates again by running through the signals through a dictionary unit. The dictionary unit is a list of meaningful words or sounds that the subject reacts to. Each word or sound has a threshold. Important units will have a low threshold, meaning that they will be selected even if the incoming signal was reduced. The thresholds are dynamic and would change overtime.

4.1.2. The Inattention Paradigm

The research performed on inattention paradigm is similar to auditory attention in that they both study perception of unselected objects. The research on inattention paradigm will present subjects with images and ask them to focus on a certain object. While the subjects are focused on the target, an unexpected object will appear for a short period of time. After the experiment is over, the subjects are asked questions on the unexpected object.

The results show similar results to the tests on auditory attention. Subjects were able to identify low level information (i.e. color, location) on the unexpected object; however, they did not perceive high level information (i.e. shape). The conclusion is that properties are either pre-attentive (can be perceived without attention) or attentive (need attention to perceive).

Many more studies followed to classify properties as pre-attentive or attentive. The original research has classified color, location, shape, and number of objects. Research performed by Mack, Tang, Tuma, Kahn, and Rock showed that textures is attentive. These research also led to the discovery of a phenomenon called inattentional blindness.

Inattentional blindness occurs when a subject fails to identify a change even though it is in their field of vision. This phenomenon occurred about 25% of the time during experiments where the subjects do not expect any changes to occur. Some surprising effects of inattentional blindness was later discovered. For example, they found that inattentional blindness actually occurs more often when the unexpected object appears in the center of the foveal fixation, rather than else where. Another finding was that inattentional blindness is linked to the meaning of the extra object. The more meaning the word or shape holds to the subject, the less often inattetional blindess will occur. For example, subjects are only blind to their own names 5% of the time; however, they are blind 35% of the time to other people's names. This was a clear indication that some form of late selection occurs because the object was first perceived, then selection was made based on the perceived object's importance.

These researches showed that both early and late selection occur in visual perception. The amount of selection changes in different situations. Evidence indicates that early selection occurs when the visual field is filled with information. Selection operates early to block and filter unnecessary information outside the focus. On the contrary, if the focus has a small amount of objects, then selection occurs late because the information outside the focus will be processed as well.

4.1.3. The Attentional Blink

The attentional blink is another phenomenon where the subjects will miss a change that occurs in their visual field. This phenomenon occurs when the image at fixation rapidly changes. Research has shown that subjects can identify all changes if the change occurs at less than 500 ms. However, if the change occurs at a rate between 200 ms and 500 ms, then there is a chance that some objects will be missed. Normally, the missing image will be ones that are shown right after a perceived image.

Attentional blink occurs because after a subject perceives an image, there is a brief period of time that the subject has ran out of attention to give. The eyes are still processing the first image when the second image appears; therefore, it chooses not to process it. Research has shown that the eyes do "see" the image; however, during late stages of attention selection, the image is ignored.

4.1.4. Change Blindness

Change blindness is another phenomenon which suggests that unattended objects are not perceived. A common change blindness experiment is as follow. Give subjects images similar to Figure 4.2, where the images are nearly identical, and ask the subjects to identify the differences. This is a surprisingly difficult task. Although, if the images were shown one after the other at the same location, the task becomes very easy. Spotting the difference is difficult because when you move your fixation from one image to the other, there is a blank interval during the process. This blank interval causes the perception of change to be difficult.

Change blindness, inattnetional blindness, and attentional blink are all indications that people on perceive and experience the things they attend to. Although there is the illusion that we can see everything in our field of vision, we actually only perceive objects and changes that we give attention to. Under normal circumstances, when we examine a given object, we will naturally examine the objects around our target, and that is the cause of the illusion.

Another interpretation for change blindness, inattentional blindness, and attentional blink is that the changes are experienced, but not remembered. This theory is named inattentional amnesia, which states that although if no attention is given to a perceived object, then the object will not be stored in our memory, hence, creating the illusion that we simply did not observe the change.

Currently, there is no definitive answer to how change blindness and other phenomena work. However, one thing for certain is that attention plays an important role in transformation our perception into something readable by our conscious and subsequently stored into memory.

So far we have studied objects that are not attended due to lack of expectation (inattentional paradigm), lack of resources (attentional blink), or some form of distraction (change blindness). Next, we will examine perception of objects that are intentionally ignored.

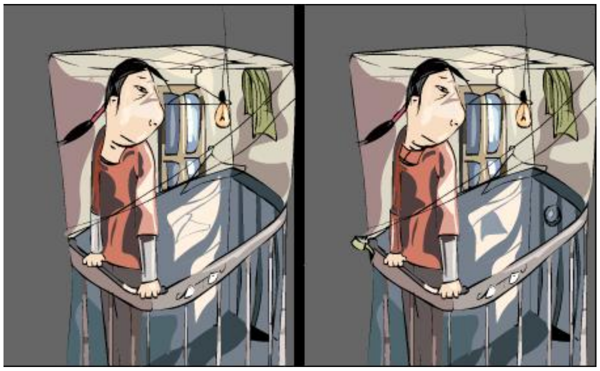

Figure 4.2

4.2. Costs and Benefits of Attention

In the previous sections, we have examined the effects of attention and how perception changes for objects that are not attended. In this section, we will examine the cost and benefits of attending to objects. We know of obvious benefits of attention such as obtaining high-level information. However, there are also costs that must be paid. We will now examine several experiments and theories that analyzes the costs and benefits of attention.

4.2.1. The Attention Cuing Paradigm

Michael Posner and his colleges at the University of Oregon put together the attentional cuing paradigm, which states that subjects respond faster to locations of interest when given a cue, but they respond slower when an object appears at an unexpected location.

The experiment is conducted as follows. Subjects must press a button when they see a flash of light. The light will either be on the left or right side of the center of their focus. An arrow sign will appear in the middle to cue the subjects where the light will appear. The light will appear 80% of the time at the correct side and will appear on the wrong side the remaining 20% of the time. The cue could also be a plus sign, which means that the chances of the light appearing on either side is equal. The three conditions of interest are as follow:

- Neutral trials – When the cue is a plus sign. The results from these trials will be compared with the other two trials to find any costs and benefits associated with attention.
- Valid trials – The light flashes at the direction given by the arrow. These trials will measure the benefits of selective attention.
- Invalid trials – The light flashes at the opposite direction given by the arrow. These trials will measure the costs of selective attention.

The results from Posner, Nissen, and Ogden's experiment shows that during a valid trial, there is an approximate 30 ms increase in response time, and during a invalid trial, there is an approximate 30 ms decrease in response time. This is a clear indication that there is both costs and benefits associated with selective attention.

Posner, Nissen, and Ogden furthered their research by varying the time between the cue and the flash of light. The results showed that it takes about 400 ms for a person to shift their attention from the cue to the location of interest. They also found that the cost of attention reaches a maximum before the benefits of attention does.

4.2.2. Voluntary versus Involuntary Shifts of Attention

There are two types of attention cues, push cues and pull cues. The cue demonstrated in the attention cuing paradigm is a push cue because your attention is pushed from the cue to the target location. It is a cue that causes a voluntary shift of attention. The other type of cue is pull cues. These cues pull our attention to them involuntarily. An example of a pull cue would be a moving object. When you stare at a still image and something suddenly moves, your eyes will automatically become attracted to that object.

There are several differences between push cues and pull cues, and they are as follows:

1. Pull cues produce benefits but not costs; however, push cues have both benefits and costs.
2. Pull cues are faster than push cues. Pull cues shift attention in about 100 ms, while push cues need 200-400 ms to complete the process.
3. Pull cues cannot be ignored. When a pull cue is presented, the subject will automatically shift their attention, thus it will not be ignored.

4.2.3. Three Components of Shifting Attention

Shifting of attention could be broken down into three components or steps. They are as follows:

1. Disengagement – Attention is removed from the object which it is currently given to.
2. Movement – Attention is moved from the previous location to the location of interest.
3. Engagement – After reaching the target, attention is given to reengage on the new object.

Each of the three components is controlled by a different part of the brain. Disengagement is controlled by the posterior parietal lobe, movements is controlled by the superior colliculus, and engagements is controlled by the pulvinar. It is possible for someone to damage a part of their brain and one of the three functions would stop. Balint's Syndrome is a condition where the patient's parietal lobe is damaged, thus they cannot disengage their attention. This topic will be discussed in details in later sections.

4.3. Theories of Spatial Attention

To understand the theories of visual attention, it is useful to have metaphors that can describe the abstract concepts of visual attention.

4.3.1. The Internal Eye Metaphor

One of the earlier metaphors developed is the internal eye metaphor. This metaphors captures two elements of attention. Attention is similar to an internal eye in that it shifts from object to object and there is a central focus where the information is processed. However, there are also several problems with this metaphor. One of the potential problems is infinite regression. If attention is like an internal eye, then this means there is an internal eye inside our real eyes. But is there also an internal eye to within the internal eye?

4.3.2. The Spotlight Metaphor

The spotlight metaphor captures the two most important concepts of attention, movement and focus. The spotlight theory states that the location where the spotlight of attention focuses is illuminated, thus the area will stand out and become easier to process. Spotlights could also be shifted from one location to the next.

Many predictions on the behavior of attention was derived from the spotlight theory. Some of the more interesting ones are as follows:

1. Rate of Motion – The amount of time needed for attention to shift is related to the distance of travel. Experiments show this prediction to be true. Attention moves at approximately 8 ms per degree of visual angle.
2. Trajectory – When attention shifts, everything along the path is illuminated. Some research supports this prediction.
3. Size – Attention has a fixed size. There is evidence that proves this prediction to be false.
4. Unitariness – Attention cannot be divided into separate regions. There are also evidence that shows this to be untrue.

Although the spotlight metaphor is popular and productive, there are also problems with this theory. One of the main problems is the size of attention. There are strong evidence that attention could change in size; therefore, some other metaphor might be better suited to describe attention.

4.3.3. The Zoom Lens Metaphor

The zoom lens metaphor is a theory that could compensate for the change in the size of attention. The core idea of the zoom lens metaphor is that the focus can zoom in to focus clearly on a small region, or zoom out to focus on a large region with less degrees of clarity. Experiments have shown that this is also true for attention. Subjects were given large objects to identify, then small objects to identify. The subjects needed time to adjust their focus (or zoom in) before they can perceive the small objects.

The zoom lens metaphor is compatible with the spotlight metaphor. Both could be used simultaneously to visual attention. The spotlight could be thought of as being able to move farther or closer to the location of interest. When it is pulled farther away, the light will be dimmed, but it can cover a larger area. When it is pushed closer, the light will be more intense, but it can only cover a small area.

4.3.4. Space-Based versus Object-Based Approaches

The metaphors described in this section assume that visual attention occurs in a given region. The alternative to this space-based theory is object-based theory, which states that visual attention select objects instead of regions.

One of the strongest evidence for object-based theory is a neurological condition named Balint's syndrome. Patients suffering from Balint's syndrome cannot perceive more than one object at a time. This will be discussed in details later.

Another evidence of object-based theory is the experiments performed by Duncan (1984). He shows subjects two similar objects. When asked about properties that both objects have, the subjects answer the question more accurately as opposed to properties that are different on each object. This shows that subjects perceive the properties of the object, as opposed to the region, since if visual attention is indeed spatial, then the properties should not affect the ability to perceive.

There are many other experiments that follow which prove the existence of object-based attention. For example, one experiment shows that attention is actually not unitary, meaning it can be split into different locations.

Currently, the experiments and theories presented are under the assumption that space-based theory and object-based theory are mutually exclusive. However, this is only true if attention only functions on one level. It is possible that attention operates at various levels of the visual system, hence, making it possible for both space-based theory and object-based theory to be true. This is a very convincing argument because there are experiments which show the existence of both theories.

4.4. Selective Attention to Properties

The theories of selective attention, such as spotlights, zoom lenses, and even object-based theories, etc., elaborate spatial selection. If visual selection is related with selection of different properties, spatial selection cannot cover all things with regard to visual selection. For example, when we try to buy a cloth, we consider different texture, color, size, shape, and so on. Thus, we need to consider different properties of an object. Can people really attend to different properties of the same object independently or does attending to one necessarily result in perceiving them all? In this section, we consider selective attention to properties. There are two experiments about selective attention to properties; the stroop effect and integral versus separable dimensions.

4.4.1. The Stroop Effect

The stroop effect is a demonstration of interference in the reaction time of a task. The psychologies, J. Ridley Stroop discovered the stoop effect in 1935. Figure 4.3 shows the stroop effect. When the color of the word is different from the color of the semantic meaning of the word, a delay occurs in the processing of the word's color and it leads to slower reaction times and an increase in mistakes.

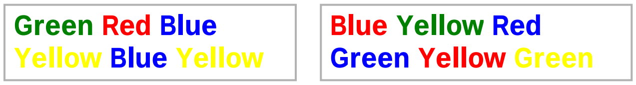

Figure 4.3. The Stroop Effect

The stoop effect is not just limited in color. Another example is that stoop effect occurs when the object name written in the object is different from the original name of the object, people experience a delay in processing.

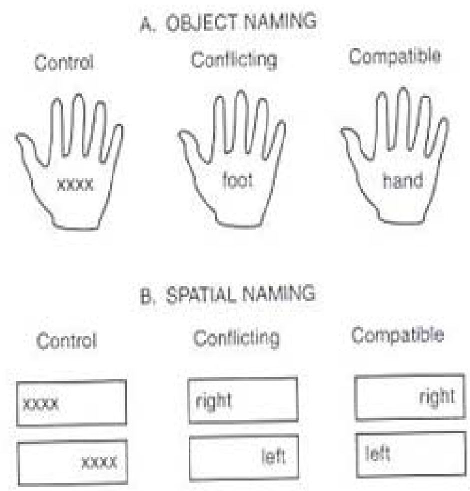

Figure 4.4. Additional Stroop Effect

In the figure 4.4, there are two types of experiments. First thing is an object naming and second thing is a spatial naming. In case of object naming, the original name of the object is different from the name written in the object. In case of spatial naming, the location of the word is different from the semantic meaning of the word. In each naming, Clark and Brownell (1975, 1976) categorize three cases which are control, conflicting, and compatible. In case of conflicting, people experience the stroop effect.

4.4.2. Integral versus Separable Dimensions

Psychologist Wendell Garner at Yale University had experiments regarding both discrete features and continuous dimensions [1]. Integrality and separability are two different relationships between pairs of properties or dimensions. Integral dimensions cannot selectively attend to one without perceiving the other. Examples are saturation and lightness of a color. Separable dimensions can selectively attend to one or the other at will without perceiving other properties. Thus, the internal representation of separable dimensions appears to be completely independent. Color and shape of an object are the examples of separable dimensions. There two dimensions seems to be processes together whenever we attend to the color of an object.

There are three conditions for determining whether the two given dimensions are integral or separable; unidimensional variation condition, correlated variation condition, and orthogonal variation condition. These conditions are defined in terms of how the unattended property varies with respect to the attended one with different blocks of trials.

Unidimensional variation condition is to classify the stimuli according to their value on one of the two dimensions while the other dimension is held constant. Correlated variation condition is to classify the stimuli according to the value on just one dimension, but the other dimension varies in a correlated way. Orthogonal variation condition is to classify according to a single specified dimension, but the other dimension varies independently (orthogonally) so that all four stimuli are presented with each block of trials. Garner examined

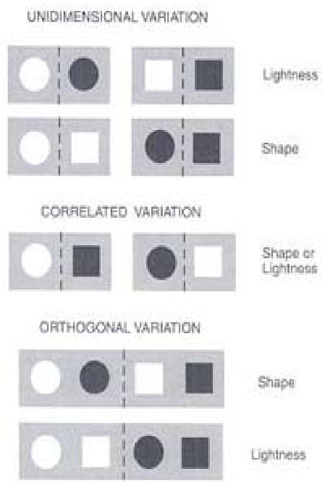

Figure 4.5. Three variation conditions.

three speed classification tasks. In this experiment, figure 4.5, people were asked to discriminate between two values of one dimension when the other dimension was held constant (unidimensional variation), when the other dimension covaried (correlated variation), or when the other dimension varied independently (orthogonal variation) [1].

Later, Garner developed two more relations; asymmetrically integral dimensions and configural dimensions. Asymmetrically integral dimensions are that the first property is separable from the second, but the second is integral with the first. This result is corresponding to the relation of color and word in the stroop effect. Configural dimensions are pairs of properties that combine to generate a new property, such as symmetry or closure. The combination of left and right-facing parentheses is an example. In this case, particular combinations generate vertically symmetrical configurations that are either closed "()" or open ")(" versus asymmetrical configurations that are left-facing "))" or right-facing "((" [1].

Garner's research shows that selective attention to different properties is possible for some pairs of dimensions but no for others. This implies that the spatial distribution of attention is not good to explain selective phenomena of perception.

4.5. Distributed versus Focused Attention

Traditionally, the visual processing before attentional selection is different from the visual processing after attentional selection. This is called as pre-attentive versus attentive (or post-attentive). But, instead of using those terms, authors in [1] introduce distributed and focused attention in terms of the mode of processing.

Distributed attention occurs when targets appear in any location. It is a parallel processing which simultaneously goes over the whole visual field. The distributed processing that occurs before focused attention occurs in parallel. Finding edges and regions defined by luminance, color, texture, motion and so forth are performed automatically and in parallel. It retrieves general information. For example, visual pop-out accounts for distributed attention.

Focused attention occurs when there is a single object to perceive. It is a serial processing which occurs after focused attention and is a sequence of attentional fixations which covers a limited region of space. Focused attention retrieves more specific and detailed information.

Visual Pop-Out

Treisman, Gelade, and Gormican (1980, 1985, and 1988) proposed visual pop-out phenomenon. Visual pop-out happens when an object stands out from others because an object is different from the others. It is detected through a distributed attention, that is, we can find a different object over the whole visual field. Visual pop-out only works when there is an extra feature.

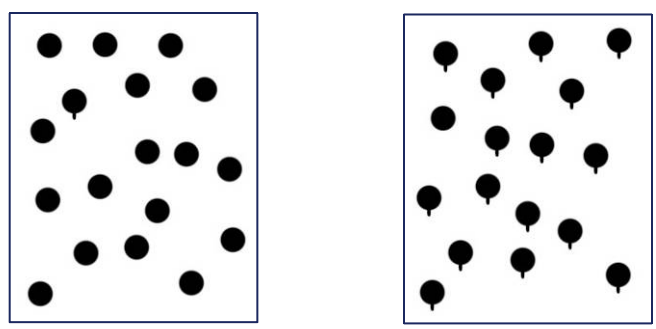

Figure 4.6 Visual Pop-Out

In the left image of figure 4.6, we can immediately distinguish the different object, since that object has distinct feature which is a circle with a line. But, in the right image, it is hard to distinguish the different object, because the target lacks feature which does not have a line in a circle. In this case, a sequential search through individual elements is required. This sequential search is regarded as focused attentional processing rather than distributed attentional processing.

4.6. Feature Integration Theory

In order to understand the feature integration theory, we first need to know binding. Biding is a process of conjoining different properties into visual object. Binding is an important theoretical problem because without some mechanism to bind properties into objects properly, an observer would not be able to tell the difference with these two pictures in figure 4.7.

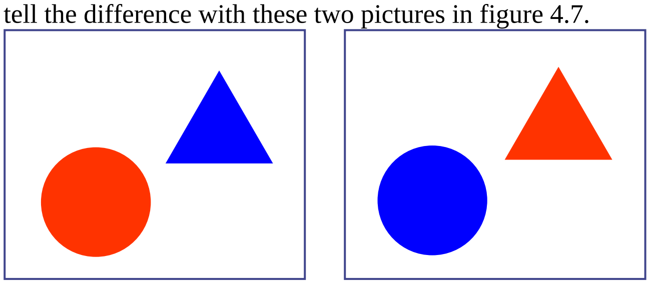

Figure 4.7. Binding

With the importance of the binding problem and a theory of focused visual attention, Anne Tresisman proposed the feature integration theory that conjoins features into objects (1980, 1998 and 1993). In other words, feature integration theory needs attention to conjoin features. In feature integration theory, features are stored in features maps at figure 4.8.

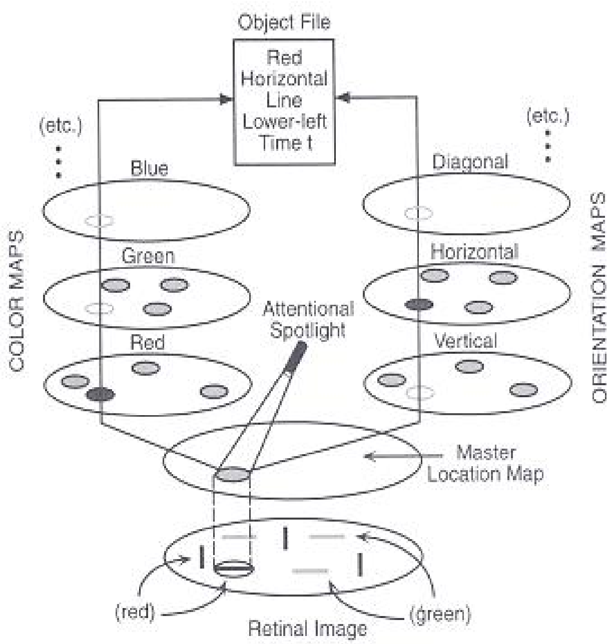

Figure 4.8. Feature Maps

In the feature map, stimuli are analyzed into feature maps which can then be accessed by a master map of locations. Focused attention selects a location in the master map and gains access to all the features at that location. There are three examples regarding feature integration theory; conjunction search, texture segregation, and illusory conjunction.

4.6.1. Conjunction Search

The conjunction search is that pop-out should occur in visual search for target objects that can be discriminated from distractor objects by an elementary feature but not for objects that require conjunctions of features (Tresisman & Gelade, 1980). In the figure 4.9 visual pop-out in A is based on the orientation. In B, the difference is color, but it is hard to distinguish the

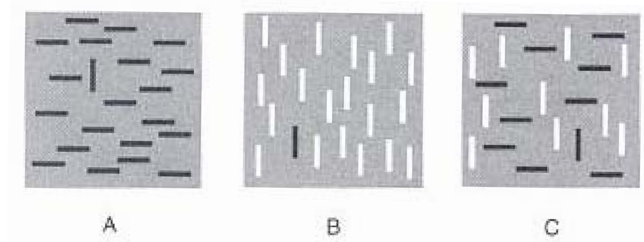

Figure 4.9. Conjunction Search

different object in C when we need a conjunction of two different features, such as color and orientation.

4.6.2. Texture Segregation

Texture segregation is that although effortless texture segregation should be possible for displays in which simple features are sufficient for the discrimination, it should not be possible for displays in which feature conjunctions define the different textures [1]. In A of figure 4.10, the stimulus is easily segregated by orientation. In B, based on color, it is easy to distinguish segregation. However, it is hard to recognize any segregation in C.

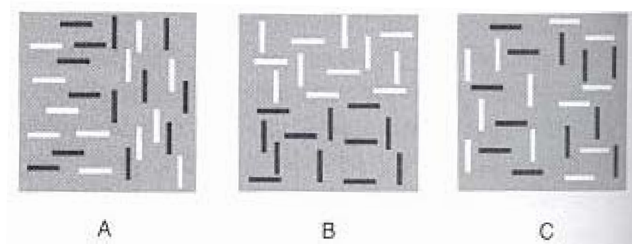

Figure 4.10. Texture Segregation

4.6.3. Illusory Conjunctions

Illusory conjunction is mistakenly merging attributes of two distinct stimuli in perception. Treisman and Schmidt (1982) tested the illusory conjunctions by flashing a 200-ms display containing a red X, a blue S, and a green T between two black digits like figure 4.11 below. It was followed by a masking display of randomly arranged colored letter fragments to eliminate any afterimage of the test display [1]. People perceived wrong parings of letters and colors. Thus, illusory conjunctions occur when short exposure and broad spatial focus prohibit attention from being directed narrowly on the individual objects in a sequence of fixations.

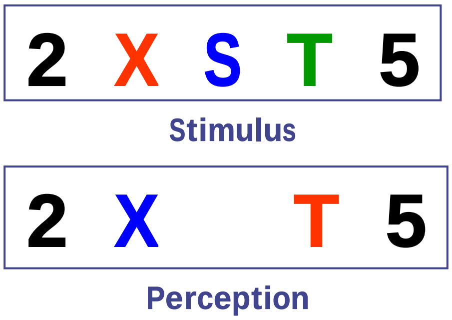

Figure 4.11. Illusory Conjunctions

4.6.4. Problems with Feature Integration Theory

After many years' experiments, Treisman found several problems in feature integration theory and modified her initial theory (1988 and 1993). There are some problems in feature integration theory. First thing is conjunction search in parallel versus serial. Since the conjunction search is based on focused attention, its processing is basically in serial. But, it is sometimes in parallel. For example, when differences in color and line length are small, conjunctions require serial search. However, when differences are large, conjunctions pop out in parallel search.

Second problem is that master location map is not coded in retinal locations. In the previous section about the feature map, figure 4.8 in the bottom part illustrates that master location map is code in the retinal locations. When we move our eyes rapidly around the array during search, this would cause the locations of the elements to shift every time. Keeping track of which items had been examined in a serial search would be extremely difficult. In terms of perceived locations, the master location map is coded in the environment rather than retinal locations.

Third problem is the pop-out of high-level features. Enns and Rensink (1990) found this problem occurring in 3D pattern. In figure 4.12, pop-out among objects appears in 3D pattern, but no in 2D pattern. If the feature maps were based on image level properties, the processing in 3D would be similar to that in 2D.

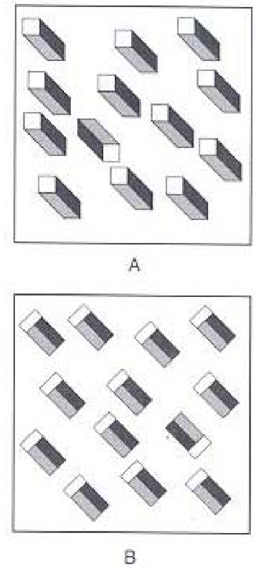

Figure 4.12. Visual pop-out in 3D pattern

4.7. The Physiology of Attention

Until now, we studied the human visual attention in terms of behavioral methods. On the contrary to behavioral methods, the idea of the physiology of attention is how the visual attention is accomplished by neural mechanisms in the brain. There are two famous experiments and other studies about the attention in terms of physiology.

4.7.1. Unilateral Neglect

The unilateral neglect is caused by brain injuries in a certain location, especially parietal lobe of the right hemisphere. In case of the unilateral neglect, a patient fails to notice objects on the opposite side of their brain injury. Figure 4.13 shows where the parietal lobe in our brain.

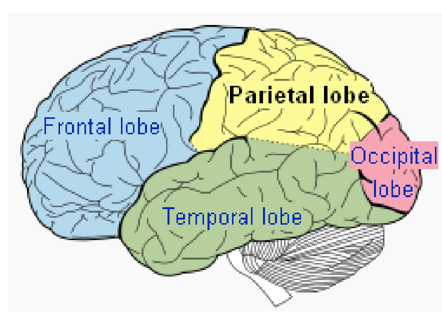

Figure 13. Our brain [2]

Figure 4.14 describes the results of experiments. When a patient with right brain damage was asked to copy a picture from the original picture, a patient only drew the half right of the original picture. He failed to recognize the left side of the image. It signifies that the patient only perceives the half of the picture.

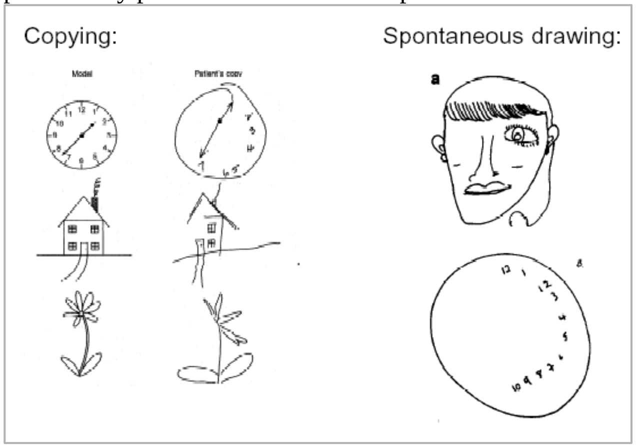

Figure 4.14.
Copying and Drawing Tests by unilateral neglect patient [3]

4.7.2. Balint's Syndrome

The Balint's syndrome is more serious than the previous unilateral neglect. It is an almost complete inability to notice anything except a single fixated visual object. Hence, patients need to close their eyes to break fixation in order to recognize others. There are four types of this syndrome; ocular apraxia, simultagnosia, spatial disorientation, and optic ataxia.
Ocular apraxia is an inability to change fixation from one object to another. Simultagnosia is an inability to perceive more than one object at a time during a single fixation. For example, even though a patient looks at a person's face, he cannot tell whether that person is wearing glasses or not. Spatial disorientation is an inability to orient and localize objects correctly including their direction and depth. Patients having spatial disorientation experience serious difficulties to perceive the visible environment and remember the places. Optic ataxia is an inability to reach out and touch an object in space.

Balint's syndrome is important in view of theoretical point since it provides neurological evidence for the existence of an object-based attentional system. Its most significant perceptual symptom is that the patient can perceive only one object at a time even when he sees more objects in the same space.

4.7.3. Brain Imaging Studies

The physiology of attention causes more studies regarding which areas of the brain are active when people (even normal people) attend something visually. Figure 4..15 shows the results of the brain imaging studies. If a person has a brain damage in the left side, the right parietal lobe can still control attention on both sides of the visual field. When there is damage in the right side of the brain, the left parietal lobe can only control attention on the right side of the visual field. This means that the brain damage in the right parietal lobe seriously affect human's visual attention rather than the damage in the left brain.

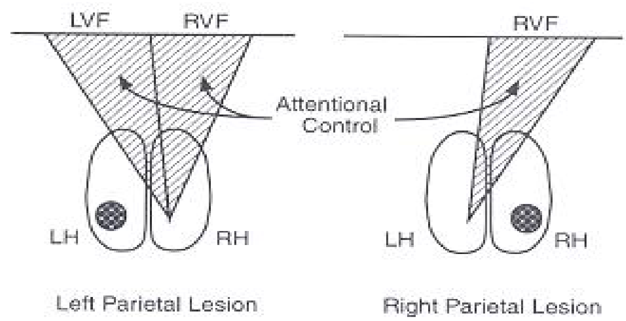

Figure 4.15.
Attentional control with left versus right parietal damage.

4.7.4. Electrophysiology Studies

Electrophysiology is a study of the electrical properties of biological cells and tissues. It mainly measures voltage change or electrical current flow in cells and tissues. Single-cell physiologists have started electrophysiology studies. In this area, only a little is known yet, but they found that selective attention to a given spatial location and object restricts the functional size of a cell's receptive field and increases its resolution for responding to specific features. As a recall, the size of cell's receptive field increased as one traces the flow of information from LGN to primary visual cortex to inferotemporal cortex where pattern recognition is thought to occur [1]. Some studies by Moran and Desimore describe that attention can influence the firing of individual cells in the visual system, but they do not identify the source of these effect. To find the source of these effect, some researchers, Desimone, Wessinger, Thomas, and Schenider (1990) examined the lateral pulvinar nucleus because it projects directly visual cortex and humans with lesions in this area have difficulty engaging attention. They find that deactivating the lateral pulvinar nucleus eliminates the gating effect of attention.

4.7.5. Pre-motor Theory

We discussed overt selection through eye movement and covert selection through attentional effects through this paper. We now consider the relation between there two selection systems. In general, we think that attention follows eye movements. For instance, we think that our eyes detect all things in the scene and then our attention focuses on something. However, it is not true. The truth is that eye movements follow attentional movements. This is the pre-motor theory. The pre-motor theory of the relation between attention and eye movement generates several predictions. In the behavioral perspective, attention will improve perception of events before the eye movement is executed, when a person executes a saccade to a particular location. Besides, if eye movement is made to a particular location, attention will occur in advance before eye movement and thus cannot be sent in some other direction. There are sound reasons regarding the pre-motor theory. Firstly, attentional shifts are purely neural operations and can be executed more rapidly than eye movement. Secondly, attention is a center process, it is better to control where the eye should be directed.

5. Conclusion

In this paper, we discuss a close relationship between eye movements and attention as mechanisms of visual selection. Among many researches with regard to the relationship between eye movement and attention, the pre-motor theory is the most dominant view. The main idea is that attention derives eye movements. In addition, attention is the major mechanism of visual selection, and the eye movement is also important but supporting task of the visual selection.