universität wien

MASTERARBEIT

Titel der Masterarbeit

Priming of visual attention in dynamic visual scenes

verfasst von

Blerim Zeqiri

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2015

Studienkennzahl lt. Studienblatt:	A 066 013
Studienrichtung lt. Studienblatt:	Masterstudium Joint Degree Programme Mei: CogSci Cognitive Science
Betreuer:	Univ.-Prof. Dr. Ulrich Ansorge

Abstract

The aim of this Master thesis is to empirically test whether the repetition of features across cinematic cuts (color, continuity, location) has an impact on visual attention. This is achieved by presenting two edited videos side by side on the same computer screens that sometimes change positions at cuts. Viewers were instructed to always look at one of the videos and ignore the other. Saccadic latency toward the target video was measured by using eye tracking equipment and served as an index of how quickly attention is reoriented after cuts. The research question is to find out whether attention is reoriented faster when more features are repeated from the pre-cut scene (continuity, color, location) to the post-cut scene in dynamic scenes.

It was expected that viewers' attention will be quickly captured by most of the features that repeat from a pre-cut scene to post-cut scene. This prediction is in line with experiments on inter-trial priming of attention, an effect that was usually studied using simplified and better controlled displays, while reflecting basically the same principles. The results confirmed our prediction that the repetition of features across the cuts in dynamic scenes accelerates saccadic eye movements from a pre-cut to post-cut scene. The insights of this study thus cannot only help to better understand the mechanisms of visual attention in more ecological viewing situations (such as film viewing), but also inform cognitive film theory and various other applications, such as developments in the field of human computer interaction.

Keywords: attention, priming, color, continuity scenes, edited videos, Bayesian surprise

Abstrakt

Ziel dieser Masterarbeit ist es empirisch zu testen, ob die Wiederholung der Merkmale entlang filmische Schnitte (Farbe, Kontinuität, Position) einen Einfluss auf die visuelle Aufmerksamkeit hat. Dies wird erreicht, indem zwei bearbeitete Videos nebeneinander auf dem gleichen Computerbildschirm präsentiert werden, die manchmal die Positionen bei Filmausschnitte ändern. Betrachter wurden angewiesen immer an einem der Videos anzuschauen und die anderen zu ignorieren. Die Sakkadenlatenz zu dem Zielvideo wurde unter Verwendung von Eye-Tracking-Geräte gemessen und diente als Index dafür, wie schnell die Aufmerksamkeit nach Filmausschnitten umorientiert. Die Forschungsfrage ist es herauszufinden, ob die Aufmerksamkeit schneller wird, wenn mehrere Merkmale sich von vor dem Schnitt (Kontinuität, Farbe, Position) auf die nach dem Schnitt Szene in dynamischen Szenen wiederholen.

Es wurde erwartet, dass die Aufmerksamkeit der Betrachter, bei den meisten Wiederholenden Merkmale von vor dem Schnitt auf die nach dem Schnitt Szenen, schneller erfasst wird. Diese Vorhersage steht im Einklang mit Experimenten an Intertrial-Priming der Aufmerksamkeit, einen Effekt, der in der Regel unter Verwendung vereinfachter und besser kontrollierte Bildschirmen untersucht wurde, während im Grunde die gleichen Grundsätze widerspiegelt. Die Ergebnisse bestätigen unsere Hypothese, dass die Wiederholung der Merkmale durch die Schnitte in dynamischen Szenen die sakkadische Augenbewegungen von vor dem Schnitt Szenen auf nach dem Schnitt Szenen beschleunigen. Die Erkenntnisse dieser Studie können nicht nur helfen, der Mechanismen der visuellen Aufmerksamkeit in ökologischer Betrachtungssituationen (wie Filmbetrachtung) besser zu Verstehen, sondern belehren auch die kognitive Filmtheorie und auch andere Anwendungsbereiche, über die Entwicklungen im Bereich der Mensch-Computer-Interaktion.

Schlagworte: attention, priming, color, continuity scenes, edited videos, Bayesian surprise

Acknowledgements

I want to thank my supervisor Prof. Dr. Ulrich Ansorge for his openness and readiness to supervise this Master thesis and his profound support toward completing the thesis. Many sincere thanks I owe to Christian Valuch MSc., for his support, good advice and patience at every turn during the whole process of completing a master thesis.

Also I would like to thank Prof. Dr. Markus Peschl and Mag. Elisabeth Zimmermann for interesting study programme as well as with organizational and administrative support.

For proof-reading in English of this work I thank very much Anja Christensen-Sylejmani.

To my family I owe a debt of gratitude for many years of encouragement and support in difficult situations.

Finally, and most importantly I would like to express my gratitude to my fiancée Ardiana Idrizi for her patience and continues support during my academic career. Also, I would like to thank her for professional as well as for emotional support.

Table of Contents

I. THEORETICAL PART

1. Introduction

In daily life situations humans are excited by a huge number of divergent visual stimuli at the same time. The cognitive mechanism that processes certain information in the cortical regions and restricts other irrelevant information is understood as attention (Cohen, Cavanagh, Chun & Nakayama, 2012). Hence, attention saves our information processing system from overload (Solso, Maclin & Maclin, 2008). Several authors agree that visual information is processed in two hierarchical stages, functionally independent from each other. First is pre-attentive stage that operates without any input limits, followed by capacity limited stage that can process one or few items at the same time. When items are transferred from the first pre-attentive stage to the second capacity limited stage of processing, they are considered to be selected items (Theeuwes, 1993). About three times per second saccades are reoriented in order to have a better projection in the retina, respectively fovea (Hollingworth & Henderson, 2002; Findlay 2005). Fovea is the most sensitive region in the retina, the one with the highest visual acuity. After a saccade is made during the visual search, the gaze rests for a while in a new location before the other saccade is done. The resting of gaze (fixation) is the process during which the visual information is encoded in memory (Hollingworth & Henderson, 2002; Smith et al., 2012). According to Kowler, Anderson, Dosher and Blaser (1995) attention and saccades are related to each other. Deubel and Schneider (1996) assume that the perceiving of visual scenes is realized by saccades (fast ballistic eye movements) and foveation (periods of fixation when the processing of information occurs). The rest of the information is restricted and lost at early stages of visual processing (Wolfe & Horowitz, 2004; Smith, Levin & Cutting, 2012). Whether the encoding process helps with the process of recognizing a location or object is discussed later on in the text.

Attention plays an important role in watching edited moving images, such as films. According to Smith et al., (2012), film scenes are comprised of thousands of images partly filmed at different times and locations. These scenes are edited together to give a viewer the impression of a continuous flow of events. Nevertheless, viewers often perceive such edited dynamic scenes as being continuous spatially, temporally and in action. Moreover, Smith et al.,

(2012) assume that perceiving a film is similar to perceiving daily life situations. While in film the scenes change across cuts, the perception of daily life situations is stable and continuous. Overcoming the sensory variation on how the film is perceived in comparison to daily life perception, and that factors such as attention and memory for visual information might play a role in this process, is one of the issues addressed in this thesis. Questions like how the visual system connects the juxtaposed shots into a continuous perceptual experience will be answered. In order to systematically investigate the cognitive processes occurring due to perception of dynamic scenes, methods such as eye tracking and fMRI (functional magnetic resonance imagery) are used by cognitive psychologists (Smith et al., 2012).

Whether roughly similar images presented in static or dynamic scenes are more efficiently recognized, is discussed in this thesis. Moreover, questions like how the viewer's integrate shots into one coherent perceptual experience, based on Gibson's ecological theory will be addressed. Furthermore, the narrative aspect of film perception and the so called Hollywood style aspect will also be addressed.

Several visual theories of attention will be taken into the consideration like the bottom up and top down factors, overt and covert orienting of attention, novelty and repetition of stimuli, edit blindness and priming effect in dynamic visual search, in order to explain the factors that play a main role on the perception of dynamic scenes.

Furthermore, the results and interpretation of results will be presented and finally there will be a summary of the master thesis.

The present work will empirically test whether the feature repetition across cinematic cuts influences the reorientation of visual attention. One question that is important to determine is how the viewer can integrate individual shots into one coherent perceptual experience. How can the viewer tell whether the post-cut image is a continuation of the pre-cut image or whether the post-cut image stems from a different location? According to Smith and Henderson (2008), continuity editing rules (repetition of the major elements from a pre-cut image to a post-cut image) play a decisive role in this process. The experimental conditions investigated in this study are conditions of color, continuity, location and interactions between them. All in all, we want to find out which features, when repeated from a pre-cut scene to a post-cut scene, mostly attracts visual attention.

The present work will contribute to the growing body of systematic research on the cognitive foundations of film perception. Besides this, priming in dynamic scenes support further research in the field of more natural scenes (ecological viewing approach). In addition, more insight is provided for film editors about the features that attract more attention after the cuts in dynamic scenes. Research of further applications including the optimization of user interfaces in websites and computer programs is provided, where feature repetition could facilitate switches from one visual scene to another, just like in film viewing.

2. Theoretical background

2.1 Cognitive film theory

Cognitive film theory differs from previous approaches such as Marxist, psychoanalytical and feminist approaches in a way that it does not propagate about historical, cultural or political factors, but rather fits with our perceptual and cognitive abilities (Anderson, 1996).

Attempting to understand nature and processes in human mind throughout different disciplines is main endeavor in cognitive science. In the same way, cognitive film theory attempts to apply research methods in film industry (Anderson, 1996). Cognitive scientists occupy themselves with topics such as acquiring new knowledge, recovering past memories, being consciously aware, etc. According to Cutting, DeLong and Nothelfer (2010), film is composed of different units and the smallest unit attended by the viewer is a shot that encourages the narrative in films.

Same questions were raised by film theorists about motion pictures e.g., shifting of attention from one scene to the other, therefore after cognitive scientists understand how a cognitive function works, film theorists apply it to the motion pictures as well. All in all, an ecological approach to film viewing attempts to take into the consideration all the issues released by cognitive science, thus it is called cognitive film theory (Anderson, 1996).

Further, Plantinga (2002) suggests that the cognitive film approach does not aim to answer every question linked with film, but its interdisciplinary approach makes it very attractive and plausible for film research. In addition, Plantinga (2002) assumes that the main aim of today’s cognitive film theory is to understand the mechanisms that are used by viewers to make sense of the scenes they watch and how they respond to those dynamic scenes.

In particular film evolved more rapidly than other forms of art within 125 years of existence, hence it is possible to explore its growth and changes that occurred (Cutting et al., 2010). The filmmakers recorded uninterrupted scenes with one single shot at the beginning of film making, usually with a single static unmoving camera. Later it was discovered that a camera could be moved, so the viewer could perceive and make sense of actions recorded from different angles. The demand for longer films increased with time, thus it increased the length of films. After a meanwhile, watching a film in a full length was very time consuming and unexciting, therefore after some time multiple shots were recorded and the cutting of the scenes was necessary (Cutting, 2005; Smith et al., 2012).

Cutting, Brunick, DeLong, Iricinschi, and Candan (2011) analyzed 160 movies that were released from 1935 to 2010 in order to find out how the Hollywood films evolved during that time. They found several differences between older and contemporary released films such as more motion, less luminance and shorter length of shots in contemporary films. In addition, they found that a shorter length of shots contained more motion in contemporary films.

Film does not only have to work for individual viewers, but it has to be reliable for all. According to Schwan and Ildirar (2010) the Hollywood style film making is easy to comprehend. They found, in their study, that inexperienced viewers were able to understand film and never confused a film clip with real life viewing situations.

However, film editors leave many scenes out of the film by using editing techniques. Nowadays, editing is basically used to link scenes with each other (Smith et al., 2012). In contrary with the static scenes, where the gaze is not fixated on the same location at the same time when perceived by multiple viewers, (Mannan, Ruddock & Wooding, 1997; as cited from Smith et al., 2012), on dynamic scenes despite the fact that the feature complexity increases, the synchrony among viewer's increases too, therefore the gaze is synchronized spatially and temporally among viewers. According to Goldstein, Chance, Hoisington and Buescher (1982) participants performed better in memory recognition of the images presented in dynamic mode, rather than stimuli presented in static mode. Editors use continuity editing rules to establish a system that tells a story as clearly as possible and give a narrative flow of continuity (Bordwell &

Thompson, 2001). According to Hochberg and Brooks (1996), memory plays a main role in perceiving a scene as continuous. After every interruption, viewers have to decide whether the current scene is continuation of the previous scene or it has different elements.

In the next part, how the film is perceived as continuous even though many cuts occur will be explained

2.1.1 Film continuity

The foremost aim of the viewer is to perceive the narrative in motion pictures. Scenes in film are perceived as continuous although they contain many cuts and spatiotemporal discrepancy across edits. This discrepancy is linear with attentional shifts which a viewer would perform when perceiving the same scene in daily life situations (Münsterberg, 1916; as cited from Smith, 2005; Smith et al., 2012). The difference between the real world and film is that the real world is perceived as continuous in time and space. Perceiving frames that contain many cuts, as a whole and integrated with each other, suggest that humans did not evolve to perceive this structure of integrity, but Hollywood films evolved to adapt our cognitive abilities to perceive films (Cutting, 2005). Moreover, Smith et al., (2012) assumes that the Hollywood style (or continuity editing rules) of filmmaking underwent congruent modifications by adapting to human attentional processes such as space, time and action continuity. Knowing how viewers succeed in dealing with differences between discontinuous scenes in films and continuous real world scenes, ensure a valuable breakthrough in cognitive processes required to perceive daily life dynamic scenes. In other words, realizing how the viewers differentiate between perceiving movies and real life scenes provides an understanding on cognitive processes used to perform this process.

Additionally on film perception, Hochberg and Brooks (1996) suggest that processing a film successfully invokes many levels of cognition. They assume that the viewer has to be able to predict the occurrence of attentional shift, in order to anticipate a mental representation of that scene. Mental representations are measurable in experimental labs and are required when viewing motion pictures. The linking of two shots with each other occurs when a question is raised from one shot and answered on the next shot.

On the other side, Gibson (1978) with ecological theory assumes that the discrepancy between how we experience visual scenes in films and real world scenes does not prevent the viewer from perceiving dynamic scenes directly. The ecological approach emerged as an alternative approach from the traditional psychological approach of empiricism. Empiricism is based on measuring simple stimuli in controlled laboratory settings, while the ecological approach assumes that humans perceive the environment directly by using real stimuli in natural conditions. The ecological approach assumes that the light strikes the visual system as an optic array, containing complex and well-structured information. In addition, humans do not construct mental images from different stimuli that enter the visual system, but rather, the perception of objects is direct. Based on research from Gibson (1978) the evidence for watching a film by integrating single shots with each other does not require any specific perceptual skills, therefore, humans perceive the content of films with a minimum of attentional effort. All in all, Gibson's ecological approach assumes that motion pictures are perceived directly, without inferences or the creation of mental states. Visual discontinuities of cuts are connected with each other based on viewer's conceptual understanding for pictured events.

Below (see sub-chapter 2.2) the phenomenon of perceiving the scenes as continued by the viewer is elaborated.

2.2 Edit blindness

The phenomenon of perceiving and experiencing films as continuous, despite many changes and interruptions occurring is labeled as edit blindness. Smith and Henderson (2008) emphasize that edit blindness is a very unique compromise that makes perceiving films as artificial products in one side and visual reality on the other side available. In a study by Smith and Henderson (2008) was investigated the relationship between edited scenes following continuity editing rules, and scenes that violated continuity editing rules was investigated. Every edit in film was distributed in one of the four conditions that followed the continuity editing rules, such as between scenes editing or low visual continuity, within scene editing or high visual continuity, match action editing or scene and action continuity and gaze match or scene and gaze continuity. The eye movements were recorded in order to find out in which case the edit blindness was higher. The participants had to detect edits while perceiving parts of some feature films. The results revealed that during the within editing scenes that followed the continuity editing rules, the edit blindness was higher, compared to edited scenes that did not follow the rules. In other words, the results found that on one side the most appropriate method to increase the edit blindness is by keeping the scene continuous as the cuts occur (within scene cuts), rather than cuts with no continuity (between scene cuts) that decrease the edit blindness. According to Hochberg and Brooks (1996) cuts between shots are used to display a long event all at once, to represent events that does not exist in particular positions or in particular times, to transmit some messages that are lost during the continuous scenes, to keep the rhythm when changes occur and to retain attention on visual interest. Moreover, Shimamura, Cohn-Sheehy and Shimamura (2014) assume that continuity is best achieved during the match action editing when the cut occurs in the middle of an action.

Furthermore, in a study by Smith and Henderson (2008) the viewers had to remember some forms of visual representation between new and old shots in order to detect a difference, but in the case of edit blindness the visual representation does not persist and the changes cannot be detected afterwards. However in static scenes, the studies have found that the viewers perceived objects with detailed visual representations and whenever an object disappeared, it was detected immediately by the viewers while they were moving their eyes on that object (Henderson & Hollingworth, 1999).

All in all, the more we understand the continuity editing rules and other processes involved while watching movies, it will deeply affect our understanding and experiences of film on one side and real world dynamic scenes on the other side.

2.3 Theories of visual search

Identification of sensory stimuli in the visual environment is crucial for adaptive behavior in daily life activities. Although a visual scene may contain many visual stimuli, attention limits the number of information by selecting only items with high priority in order to visually process further. Selective visual attention is investigated very often via visual search paradigms where the viewer has to detect a particular target among many visual distractors. According to Becker and Horstmann (2009) the importance of attention in perception has been researched through visual search paradigms.

In a study by Treisman and Gelade (1980) they proposed a feature integration theory of visual attention described in two phases. They assume that in the first pre-attentive phase the viewer is focused on single feature that is distinguished from other features backed up by basic independent features such as color, size, orientation, edges, movement etc. For this single distinguishable feature, the viewer does not spend any effort on being attentive, since it pops-out automatically. The search for the first pre-attentive phase is an efficient or parallel search across the display and it is not influenced by the number of surrounding distractors or total number of items in the display. Obviously, in this phase participants detect the objects based on their parallel processing. In the second attentive phase the target does not pop-out automatically, so the objects are identified after the attention is focused on them or serial search is applied. In serial search, the conjunction of single features are integrated together in order to identify the object. With an increased number of objects on the display, it increases the searching time for the target. Nevertheless, both processes in normal conditions are integrated with each other. The initial automatic coding in the whole visual field are basic features, and the objects are identified in a later stage when the attention is focused on them. In some specific conditions, the parallel and serial search processes can operate independently from each other. On the other side, without focused attention features are not matched with each other and the object cannot be perceived as complete.

Wolfe (1994) on the other side proposes a Guided Search model aiming to explain human visual capability to find a target. He suggests that the initial stage of perceiving visual objects is the pre-attentive or parallel processing of basic features. The processing of these basic features guides the deployment of attention in the later stage by triggering the establishment of independent maps for different colors, locations and so on, called feature maps. The feature maps rush our visual field and the more activation in certain locations, the higher is the probability that the attention will be directed in that location. At a later stage the most interesting locations or combinations of different basic features for further processing are spotted. On the whole, there are two principles of how the attention is activated. The first is stimulus- driven or bottom up activation where the strength of activation is not dependent on the viewer's previous knowledge, but rather the degree at which it is differentiated from other stimuli. When target features become more similar with distractor features, the target becomes difficult to detect and therefore the search for the target will increase. In the cases when more distractors share common features and become more similar with each other, the search for the distractors will decrease, and the search for the target will increase, therefore it becomes easy to detect. The second principle is the top down activation or limited capacity phase where the attention is guided toward desired object based on previous parallel processing. In the second phase more complex processes are processed like face recognition, reading, object identification etc. All in all, the weighted sum of all top down and bottom up activation determines which features are appropriate in order to establish the relevant activation map.

2.3.1 Top down and bottom up influences on visual selection

In daily life situations, perception may seem very simple with just a little effort, but in reality it is a very complex process. How perception is related to cognition and whether cognition is directed from what we perceive are some questions leading to the explanation of information processing.

The incoming visual information and previous experiments on visual research, has led psychologists to differentiate between two different ways of directing attention towards new sensory stimuli present in the environment. First arguments support the idea that attention is directed based on top down orientation such as goals, past experiences, intentions etc. When lacking top down orientation, attention is relocated on the features that distinguish one item from surrounded items, respectively the most highly salient target in the visual area that pop–out on our retina (Theeuwes, 2010). Salience is described as a topographic feature map consisting of mostly bottom up features such as color, intensity and orientation maps (Itti, Koch, & Niebur, 1998). The higher the saliency on an object, the higher is the chance that the fixation will be directed on that position (Itti & Koch, 2000). Bottom up processing refers to an automatic process governed by the characteristics of the perceptual input alone or low level features such as color, edge orientation, motion and luminance. Further, Theeuwes (2010) suggests that after the attention is oriented toward highly salient objects, the process of selecting follows and then the process of identification. Once, the selected object does not look similar to the target, the immediate detachment from that location follows without any effort. On the other side, when the selected object looks very similar to the target, the detachment is slow and effortful.

In a study by Mital, Smith, Hill and Henderson (2010) the role of low and mid-level visual features (motion, color, luminance, edge orientation) was investigated during the dynamic scene viewing and the influence of these bottom up features in gaze allocation. Results showed that low level visual features have a little visual effect on gaze allocation in static scenes, but including the motion in dynamic scenes led to significantly higher level of visual effects in gaze allocation and more alignment among viewers. All together the viewers attended and almost at the same time fixed their eyes on the same moving objects. The phenomenon of viewers gaze synchronization while perceiving moving objects in dynamic scenes, is called attentional synchrony (Smith, 2013).

Further, Theeuwes (2010) found that during the first phase of information processing (below 150ms) the visual selection is done based on low level features or bottom up processing. On the other side in a later phase (above 150ms), the information processing is done volitionally or based on top down processing. Top-down processing refers to explicit instructions given to the observer and is based on higher cognitive factors such as expectations, goals, and memories. In other words, the viewer chooses his/her selections from the environment volitionally (Lachter, Forster & Ruthruff, 2004; Theeuwes, 2010; Wolfe, 1994; Wolfe, Butcher, Lee & Hyle 2003). Wolfe et al., (2003) investigated whether viewers are more influenced by top down or bottom up information during the visual search. In order to minimize the influence between top down and bottom up information, conditions in the experiment were e.g., blocked, mixed, fully mixed etc. The blocked condition was highly filled with top down information, i.e., the target had a single and consistent feature. Whereas, in a fully mixed condition the target was based more on bottom up information and lacks top down information, i.e., colors are randomly changed and participants cannot guess what color shows up next. Results show that the reaction times were slowest when the condition was fully mixed respectively the trials were pretty much isolated from top down guidance, and the reaction times were the fastest in trials with blocked conditions when top down information was the highest.

On the other side a study by Frank, Vul and Johnson (2009) assume that top down factors, when relevant or have a cognitive semantics to the viewer, play a main role in gaze allocation, rather than bottom up features. The researchers recorded eye movement of infants and adults in order to see visual preferences while watching dynamic scenes (animated movies). They found that infants, as young as they were, had tendency to saccade along bottom up features, but as they grow up it increased the tendency to focus more their attention at locations with more cognitive semantics (in this experiment e.g., faces).

Besides bottom up and top down orientation of attention, stimuli are oriented by overt and covert shifts of attention. This is a topic in sub-chapter 2.3.2.

2.3.2 Overt vs. covert attentional shift

Attention and eye movements are linked very much in daily life situations. While looking at visual scenes, humans shift their gaze consciously across the scenes and perform saccades that adjust the retinal projection at the fovea toward regions in the visual scene that we are interested to process. Such orienting of attention that involves saccadic eye movements in order to identify objects and analyze object spatial relations is called overt orientation (Itti & Koch, 2000). In the cases when the eyes are kept still, it is the covert orienting of attention that takes place (Rizzolatti, Riggio, Dascola & Umiltá, 1987). Covert search is orienting of attention toward regions of interest in visual scenes while the gaze is fixated at one particular location, respectively attention is oriented while eyes stay at rest (Findlay, 2005; Lachter et al., 2004; Nothdurft, Pigarev & Kastner, 2009). According to Posner (1980) overt orienting of attention can be measured by eye and head movements while covert attention can be measured by an internal central mechanism. Covert attention is more usual at laboratory settings, but there are some social situations as well, such as, for instance, when eyes moving would mean revealing a secret. In a study by Klein (1980) whether attention and eye movements are related to each other was investigated. In this study it was hypothesized that attentional shifts and saccadic eye movements share a common oculomotor mechanism. The oculomotor readiness mechanism reveals the evidence of a relationship between attention and eye movement proposed via two assumptions:

The first assumption is the readiness of the observer to saccade at a certain desired location that will facilitate detection of stimulus emerging on that same location, rather than when stimuli emerged on a different location.

The second assumption is having a cognitive control of visual attention accomplished by only preparing but not moving the eyes to a certain location which has to be attended.

Dual-task paradigm was used for this experiment. In saccadic response task, participants were instructed to remove their eyes from a primary central target to the peripheral target immediately when they saw an asterisk. It is assumed that the procedure is simple and will urge the participant to prepare saccade in advance. The acquired data did not support Klein's oculomotor hypothesis, so he rejected the assumption that stimulus detection is accelerated when the eye movements are prepared. On the other side covert attention is accomplished, despite the oculomotor mechanism (without executing saccades).

Rizzolatti et al., (1987) opposed to Klein's rejection of oculomotor readiness mechanism by providing more data on supporting their reformulated Premotor theory of attention. They suggest that the shifting of attention toward stimuli is preceded by eye movements on that location. In this study participants were instructed to respond manually as fast as possible when they perceived a visual stimulus. The participants were cued to attend different possible locations for the visual stimulus and they had to respond despite target appearance to the cued or not cued location. The results showed that the proper orientation of attention provided a small but significant aid in acceleration of reaction times, while improper orientation of attention provided a high delay in time required to respond manually to the visual stimulus. Further, the delay in reaction time was increased depending on the difference between the attended location and the stimulated location. Moreover, additional time was needed to react when the stimulated and attended locations were set on opposite sides. Based on results, Rizzolatti et al., (1987) suggests that overt and covert attention share the same mechanism for orienting attention as a central programming. In addition, they propose that the motor system prepares eye saccades for visual attention; therefore they are much related to each other. Another study by Nothdurft et al., (2009), with humans and monkeys, investigated whether the reaction time was shorter during the covert (fixed gaze search) or overt (free gaze search) attention. Results show that both humans and monkeys were faster in reaction times due to covert search rather than due to overt search. According to Findlay (2005), overt and covert attentions are seen as complementary to each other.

In the case of covert attention where the orienting of attention is done without eye movement, the principle of peripheral inhibition is applied, while leaving common central programming for orienting of attention as unchanged (Rizzolati et al., 1987).

In addition, Deubel and Schneider (1996) propose that the oculomotor process and covert attention are strongly linked with each other. They used a dual - task paradigm to measure the interaction between voluntary oculomotor saccades and covert shifts of visual attention. Subjects were instructed to fixate their gaze on a cross in the middle of the screen monitor. On both sides of the screen, there were numbers inside the colored ellipses, where number 1 was inside the red ellipse, number 2 was inside the green ellipse and number 3 was inside the blue ellipse. After some time of the gaze fixation in the middle of the screen the cross reversed to array, orienting participants to the location where they should saccade. Based on the array color, the participant has to saccade on the similar color ellipse, i.e. if blue array, the saccade has to be directed at the third ellipse on the right side. Participants were instructed to hold on the saccades, and to release them only after the color ellipses are left in the monitor, but not the colored array nor the numbers inside the ellipse. After this the participant has to manifest the identity of the discrimination target. The same paradigm was applied to the second experiment, except that the participants knew the location in advance, where the discrimination target would appear. Both experimental results revealed that the discrimination performance was the best when the saccadic and discrimination target were in the same location, and otherwise the discrimination performance was the slowest when the saccadic and discrimination target did not coincide with each other. The results found that covert attention and voluntary oculomotor saccades are not independent, but in contrary, they are highly dependent on each other. In the next section, different approaches of attracting human visual attention are examined.

2.3.3 Novelty vs. repetition

Theories on how the visual stimulus is oriented during the perception of dynamic natural visual scenes are divided into two main groups. The first model supports the assumption that attention is oriented on novel stimuli or previously not attended visual stimuli, while the second model does not support the assumption that novel stimuli plays the main role. Based on newly developed methods of data analysis and the Bayesian framework of surprise that quantifies what stimuli attracts the human attention in dynamic scenes mostly, Itti and Baldi (2009) assume that humans are mostly reliable in saccading and fixating surprise or novel stimuli. There are two main elements defining surprise: the first element is that the surprise occurs when uncertainty exists, such as intrinsic stochasticity, missing information or limited computing resources; the second element is subjective expectancy of the viewer. Different viewers perceive the same information with a different amount of surprise, and the same viewers perceive the same information differently when measured in different times (Itti & Baldi, 2009). In addition, they suggest that the novelty theory framework can incorporate bottom up and top down information together. Hence, the combination of already known information from previous knowledge can be incorporated with new incoming information. In an eye tracking study by Itti and Baldi (2009), it was hypothesized “that surprise attracts human attention and gaze in natural scenes”, whereas viewers were watching video clips in different environments (daytime, nighttime, video games, sports etc.). They found that stimuli with increased motion, salience and surprise, attracted the most viewers’ attention during the experiment. Saccades were attracted in more than 72% of the cases toward surprising locations while watching videos (Itti & Baldi, 2009).

A second group of authors such as Maljkovic and Nakayama (1994) oppose to the assumption that novel stimuli attracts attention. A visual search experiment found that attention is reoriented toward similar visual content. When viewers anticipate the color of a particular target then the processing of that feature will be faster. This occurs due to short- term memory encoding of specific details in a visual scene, which are memorized during the fixation process. Gegenfurtner and Rieger (2000) found that color contributes to recognition in two different ways. At the early stage of visual processing, participants performed better at images presented with color information for 16 msec, even though they were tested in black and white. Participants could not detect a color in the target image explicitly, but they were able to identify the structure and form defined by color. In a later stage of visual processing, when the color in the image was presented longer, the recognition was better compared to black and white images, only when the testing was performed in color. They assume that the strength in color is due to the representation of color in short-term memory. All in all, the processing of color was faster when it was repeated in a successive trial, meaning that the color feature is primed by the short-term memorization of the same color, and the improved performance in finding a target in a later stage of visual processing.

According to Hollingworth and Henderson (2002) detailed visual information is maintained in memory during the previously fixated objects. Results from this experiment reveal that participants were able to detect changes that occurred in previously attended visual objects, which was not attended when the change occurred. In other words, fixations of specific locations or objects support later memorization of the static scene. Another study by Valuch, Becker and Ansorge (2013) in static scenes was found that participants were more eager to perceive previously fixated locations during the process of recognition, but the experiment did not guarantee that successful recognition is based on previous encoding. Moreover, Valuch, Ansorge, Buchinger, Patrone and Scherzer (2014) in an eye tracking experiment investigated whether viewers reorient attention faster in dynamic scenes where most of the visual content is similar from a pre-cut to a post-cut scene, versus scenes where most of the visual content in post-cut scene is different from a pre-cut scene. Saccadic reaction times (SRTs) were measured. Viewers had to pay attention to a relevant video and to ignore an irrelevant video. Both videos were shown next to each other. In contrary with Bayesian surprise, Valuch et al., (2014) found that SRTs were shorter after the cuts where most of the features from a pre-cut to a post-cut scene were repeated. In other words, attention in dynamic scenes is best captured by the scenes that repeat most of the visual content from a pre-cut to a post-cut scene. Becker and Horstmann (2009) found that when the target and distractor repeated the features from one trial to the next trial, the visual selection was faster.

2.3.4 Priming effect in visual search

Priming is a modified representation of a feature which enhances faster processing and facilitation on searching of the same feature. Maljkovic and Nakayama (1994) found that the priming process lasts about 30 s and pop-out priming is an automatic process that does not require any conscious effort from the observer to control it. In addition, pop-out priming occurs due to implicit short-term memory. The established memory trace for subsequent trials facilitates the processing of the same color trials in a passive way. If primed feature associates with the target then it is found easily (Kristjánsson et al., 2002). Several studies (Born, Ansorge & Kerzel, 2012; Kristjánsson, Wang, & Nakayama, 2002), found that in visual search tasks, viewers react much faster when the same target color is repeated from a previous trial (bottom up orientation), rather than a different color**.** Moreover, Wolfe and Horowitz (2004) in their study assume that some basic features of visual stimuli such as color, size and orientation can guide attention in the path from the early vision stage to the object recognition stage.

According to Theeuwes (2010) when viewers in experimental settings receive an information cue at which side the target will show up, they are faster and more accurate in finding a target. According to Hoffman (1975) participants' reaction times were decreased and the response accuracy was increased when an indicator was presented prior to stimulus display, indicating the presence of stimulus on that location. In addition, Posner (1980) with his paradigm on spatial cueing task found that measurements of covert attention are as accurate as measurements of overt attention. In this study the effect of covert orientation of attention as a result of different cue conditions was investigated. Participants were tested for how long they could ignore the irrelevant stimuli, while they were instructed to fixate their eye gaze at the plus sign or an arrow in the middle of the monitor that contained two boxes parallel to each other. The cues that indicate the possible relevant target stimuli were presented to participants. For a short moment, in every single trial the cue appeared in one of the boxes and then was removed. After the cue removal, the target reappeared in one of the boxes. The participants had to report the target immediately after they detected it. Reaction times were measured. The results demonstrated that participants performed faster and more accurately in trials when they were cued in advance, whereas they reacted slower and less accurately when they were not cued. All in all, the results showed that cuing orients visual attention, while the eye gaze remains fixated.

In a pop-out priming paradigm by Maljkovic and Nakayama (1994), viewers were primed while they were searching for random pop-out green diamonds in a black background. Pop-out search experiments help researchers find out the effects of feature priming while searching for the target. The results revealed that the search for the target was fastest when both the colors from the target and distractor were repeated from the previous trial, or appeared in the same location as the previous trials and vice versa. In addition they suggest that priming in pop-out search is autonomous and there is no voluntary control.

In another pop out search study that supports repetition of visual content from a pre-cut scene it was assumed that the repetition of the target and non-target features across trials, facilitate the search for the target, respectively it shortens SRTs (Becker, 2008a). Conversely, saccadic latencies should be prolonged if features are not repeated from a pre-cut to a post-cut scene (Becker, 2008b).

Besides pop-out search, the priming effect is applied in conjunction with search paradigms where two or more basic features are conjoined. Kristjánsson et al., (2002) assumes that the priming effect plays the major role during the conjunctive visual search. In their paradigm they used the conjunction of two features i.e. horizontal vs. vertical and red vs. green. Results showed that priming played a major role in the conjunction search. Reaction times were faster when the target and distractor repeated features from a previous trial. In addition, they found no difference in facilitating the search for the target whether the observer knew in advance where the target was or not.

In their study Treisman and Gelade (1980) were interested to find out whether the participants could change the processing from conjunction search (green T) to parallel processing with a lot of practice. Even after the practice, the participants showed that the change in processing (conjunction to parallel) was difficult and this unification of color and shape was not reached.

In an eye tracking study, Becker and Horstmann (2009) investigated whether priming has an impact on search efficiency, when the target and distractor features were repeated from one trial to the next during conjunctive search. Participants had to search for conjunction of target features among distractor features. The conjunction features for the target and distractor altered across trials in the display between different colors (red and green) and orientations (horizontal and vertical). The results reveal that repeated features of the target and distractor over the trials had a significant positive effect on conjunctive search for the target. When the target and distractor features were repeated, search efficiency increased. In addition, they found that the amount of the objects in the display did not play any role in an effective search.

Becker (2008b) investigated whether the feature priming effect is driven by features in the attentional stage to select the target or whether priming helps during the decisional stage. In this study, participants had to find a target item (black square) that could be either smaller or bigger than the other items. The results revealed that priming accelerated the search in the early attentional stage, which suggests that repeated features reduced the searching times and facilitated finding a target. On the other side the priming effects on the decisional stage could not be confirmed.

Studies of priming effect assume that low level features such as color are encoded separately and independently from the distractor. They assume that visual selection depends on the context. Becker, Valuch and Ansorge (2014) measured viewers' eye movements with a pop-out search paradigm where same colors were repeated among trials or different colors were repeated in different trials. They found that elementary features such as color are not encoded separately, but color priming is based on the context.

II. EMPIRICAL PART

3. Outline of the research question: Priming of visual attention in edited videos

This study aims to empirically test whether attention after cuts in edited videos is reoriented more quickly when features are repeated from a pre-cut scene to a post-cut scene across cinematic cuts. It is very important to determine how the viewer realizes whether the post-cut image is a continuation from the pre-cut image or whether the post-cut image stems from a different film scene. This strategy was predicted on the basis of the time the subjects spend on scanning the post-cut image for repeated features to arrive at the most basic decision about whether or not the post-cut image is a continuation of the pre-cut image. A tutorial review by Kristjánsson and Campana (2010) support this assumption by finding that the search for the target was facilitated and an overall better performance was shown when more target features such as color, shape and same spatial positions were repeated across trials, rather than in trials when the target changed the features and the spatial position. Moreover, Gegenfurtner and Rieger (2000) provide evidence of color contribution on memory encoding and memory representation. In our study this evidence is demonstrated when participants perform better in trials when the color condition is repeated from a pre-cut to a post-cut scene, rather than in trials when the color is absent in pre-cut and present in post-cut scenes.

In contrary with Itti and Baldi (2009) framework, stating that stimuli with increased motion, salience and surprise, attracted most viewers attention during the perception of videos in dynamic scenes, this study assumes that feature repetition plays the main role in attracting human attention. In comparison to searching for and orienting towards novel features in the post-cut image, the decision about whether the actual image is a continuation of the pre-cut image can be achieved at some point before the post-cut image has been exhaustively scanned for its content, therefore, we assume that the repeated features play the main role in this process. The first evidence supporting this hypothesis was presented in a short paper by Valuch, Ansorge, Buchinger, Patrone and Scherzer (2014). The present study builds on these previous results using an improved study design and a better controlled procedure.

Feature priming plays a role too. Based on the priming literature, the prediction is that saccadic latencies to the target video after cuts will be shortened due to a repetition of features from a pre-cut scene to post-cut scene (Becker, 2008b). Conversely, saccadic latencies should be prolonged if features are not repeated from a pre-cut scene to post-cut scene.

Finally, based on the results from previous studies in static scenes we expect similar results in dynamic scenes, in that saccades to the target movie will be faster in scenes that have a high feature similarity between the pre-cut scene and the post-cut scene.

4. Method

4.1 Participants

Thirty-six participants (25 females) volunteered from an internal database at the Faculty of Psychology, University of Vienna. Average age was 22.7. Participants were naive with respect to the experimental hypotheses, and received partial course credit for participation. Participants had normal or corrected-to-normal sight and informed consent was signed prior to the start of the experiment. In order to exclude participants with color blindness the Ishihara color test was used prior to the session. In addition the dominant eye was determined using the Porta-test prior to the beginning of the test.

4.2 Apparatus

Eye movements were monitored by EyeLink 1000 Desktop Mount Eye Tracker (SR Research Ltd., Kanata, Ontario, Canada), operating at sampling rate of 1000 Hz. By using a standard 5-point calibration, the dominant eye of each participant was calibrated via eye tracker. Drift check was applied at the beginning of every 20 blocks, in order to reassure the system calibration was working. Saccade latency was measured after each trial when the video switched position with very high accuracy. Participants viewed the experimental stimuli programmed in Matlab (MathWorks, Natick, MA, USA), using Psychophysics and Eyelink toolboxes (Cornelissen, Peters & Palmer, 2002). Stimuli were presented on a Multiscan G400 (Sony, Inc.) 19-in.color CRT monitor, with resolution of 1,280 × 1,024 pixels at 60 Hz. The experiment was run on a personal computer (Microsoft Windows XP (SP3) with an Athlon Dual Core CPU (AMD Inc.), a GeForce GT220 (Nvidia, Inc.) and 2,048 Megabytes of RAM.

4.3 Stimuli

As stimulus material twenty blocks of silent sport videos with 456 cuts were used, preselected from various DVDs at the Vienna public library and browsing online videos. Each cut equals one trial in the experiment. Stimuli were ten sets of two movies (e.g., surfing vs. skiing) that were presented simultaneously at the same time in a monitor, on the black background. Cuts always occurred simultaneously in both videos. The video repeated the same sport throughout one block. In every trial presented, one of the movies was the target movie and the other was the distractor movie. Both, target movie and distractor movie, contained simultaneous cuts along the scenes, but were different in content. All re-edited videos were encoded as silent tracks, with 25 frames per second and a resolution of 400 × 300 pixels. The participants saw the videos as 9.7 × 7.1° because the viewing distance was held constant (64 cm) from the monitor during the experiment by using a forehead and chin rest. The videos were counterbalanced in the center of the screen by 7.2° towards left or right, and 31 × 24.2° vertically centered in full-screen.

4.4 Experimental design

Three types of combinations were used in a factorial design of Continuity. The factor of Continuity was manipulated across three conditions whether the movie appeared the same or different after the interruption. In the No Continuity condition the visual content was missing from a pre-cut scene after the interruption, but the main idea (semantics) was similar. In the Continuity condition the visual content was repeated from a pre-cut scene after the interruption, but it was shown from a different perspective. Finally, in the Full Continuity condition the maximal visual continuity was shown after the cut occurred. The cut was placed during the ongoing scene and repeated all features after the interruption.

The factor of Color was manipulated across four conditions contained four conditions based on manipulation of color before and after the interruption. The four Color conditions are: when the scene was shown in color before the interruption and continued in color after the interruption (Col-Col); when the scene was shown in Color before the interruption and continued in Black and white after the interruption (Col-BW); when the scene was shown Black and white before the interruption and appeared in Color after the interruption (BW-Col) and finally; when the Color scene reappeared in Color after the interruption (Col-Col). Finally the factor of Location was manipulated by switching sides of the appearance of the video (Location same and Location different). All conditions between three factors of Continuity and four factors of Color were fully crossed and were measured among subjects. Figures 2, 3 and 4 illustrate the combination of different conditions of continuity, location and color.

Fig.1 presents a setup of a single block during the experiment.

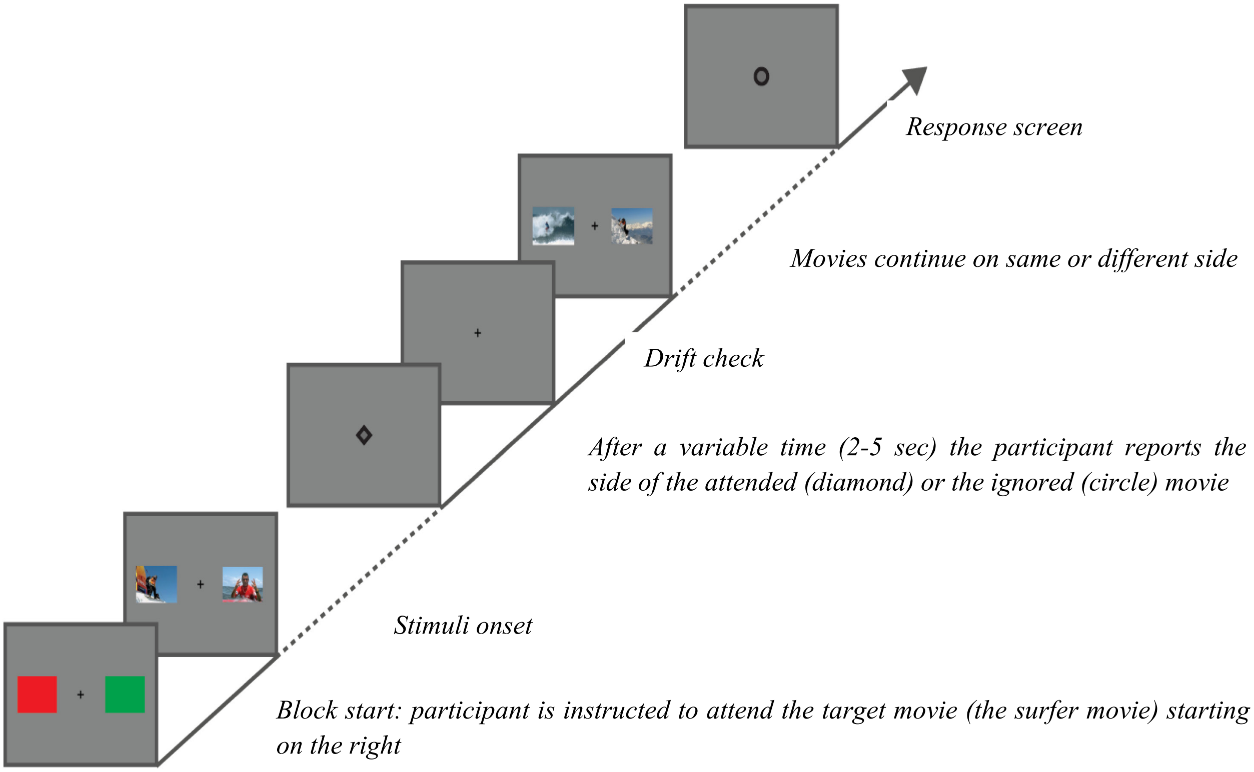

Figure 1. Setup of a single block during the experimental procedure.

Figure 1 shows a setup of a single block of the experimental procedure. The green box represents the task relevant video that has to be focused on, while the task irrelevant video in the red box has to be ignored. The videos switched sides simultaneously or reappeared at the same location after each cut. Parallel with the side change, the participants moved their gaze to the other side as well. The experimental procedure continues until the end of the block. Altogether 10 blocks like this were played in this order of around 2 min in length. The figure was adapted from Ansorge,Valuch and König (2012).

Figure 2. Location change after the cut.

This figure presents the scene with the location change after the cut occurred. Although the post-cut scene is Continuous, the location from a pre-cut to a post-cut scene was changed. The other combination in the first factorial design is on the same side of the video after the cut.

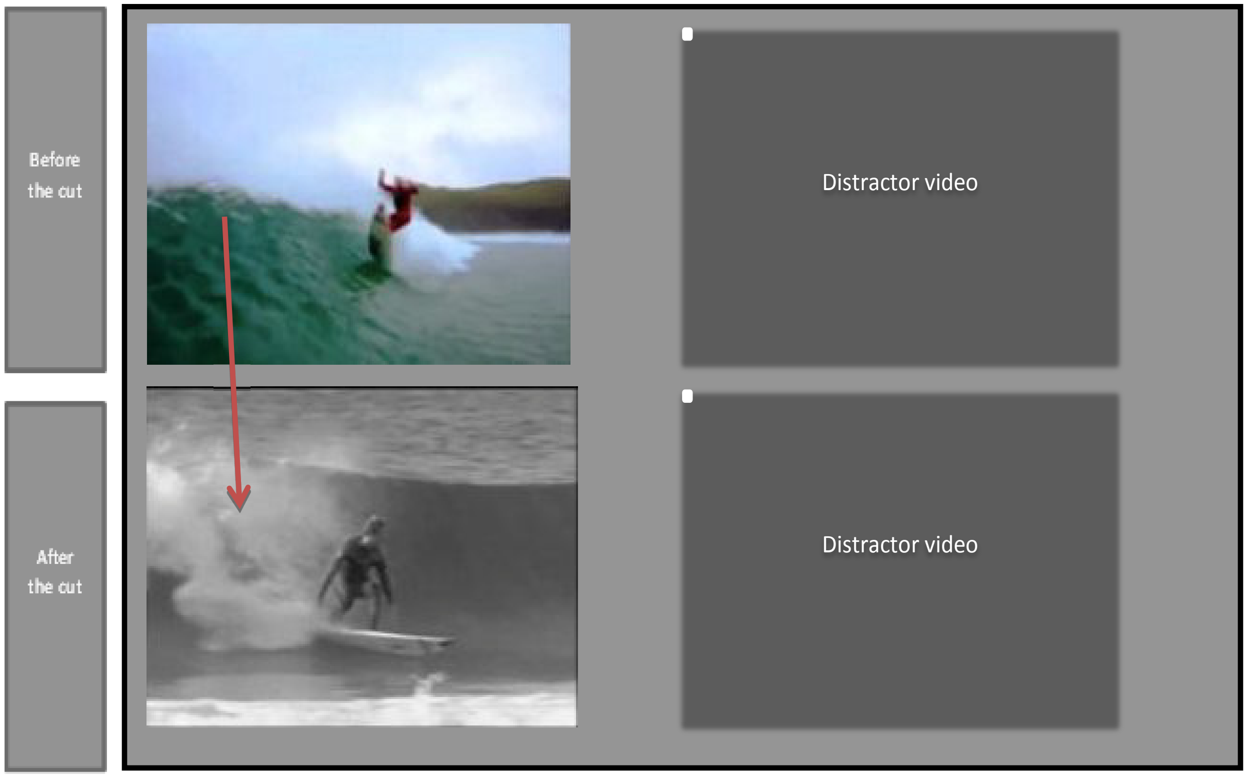

Figure 3. Color to black and white change after the cut.

Figure 3 shows changes from color to black and white (Col-BW). Most of the features were repeated from a pre-cut scene to post-cut scene besides the color feature. Color switched to black and white after the interruption. Other conditions within the factor of Color are: Color is repeated from a pre-cut to a post-cut scene (Col-Col); Black and white is repeated from a pre-cut to a post-cut scene (BW-BW); Black and white is changed to color from a pre-cut to a post-cut scene (BW-Col).

Figure 4. Continuity change after the cut.

Figure 4 shows the change in continuity from a pre-cut to a post-cut scene (No Continuity). The other two combinations within the condition of Continuity are: In the case of Continuity condition the major features are repeated from a pre-cut to a post-cut scene (Continuity), and in the condition of Full Continuity condition, all features from a pre-cut are repeated to a post-cut scene (Full Continuity).

4.5 Procedure

After the completion of the screening tests, participants were sitting comfortably in front of the computer. Instructions on how to perform the experimental tasks were displayed on the monitor. Participants understood that a set of silent sport movies would appear during the 20 blocks with 456 trials recorded as eye movement data. For each trial two movies were shown simultaneously and instructions are given to the participants to keep their eyes on the task-relevant movie and ignore the distractor movie as much as possible. In each block, the target movie was announced with a green box and the distractor movie was announced with a red box. For example, if a soccer video and a basketball video were presented, the participants were instructed (by a green instruction cue) to watch and follow the soccer video. Throughout the experiment, a pair of movies reappeared two times (the task relevant movie and the irrelevant movie). In the first part of the experiment, the participants focused on ten blocks of videos by focusing on task relevant videos and ignoring irrelevant videos. In the second part of the experiment, the same blocks of videos reappeared, with the condition that the participants had to ignore the videos they were focusing on in the first part, and they were focusing on the videos they had ignored in the first part. In other words, every block with two videos appeared twice, so the videos that appeared as the task relevant videos in the first part, switched to irrelevant videos in the second part. After each trial an interruption occurred. Participants were instructed to follow the task relevant movie immediately when it was shown after the interruption. If the location changed, participants had to move their eyes to the new location by shifting their visual attention, while the SRTs were measured with eye tracking equipment. Based on Deubel and Schneider (1996), SRT is a veridical and sensitive index of attention.

In order for participants to be attentive and to place their gaze in the middle of the screen, after each interruption a filler task was introduced. Participants were instructed to press a button after each trial (in a variable response screen), where participants had to report the position of the target movie according to a variable stimulus-response mapping; e.g. if the target movie reappeared at the same side they had to press key # 8, for the other side the participants had to press key # 2. Participants were encouraged to respond as quickly as possible while maintaining high accuracy.

5. Results

5.1 Data

Since the process of attentional selection and saccades are strongly related to each other, we analyzed the properties that orient attention after each cut. A total of 16,416 trials of eye movement data were recorded. A point of interest in this experiment was to find out how the attention is reoriented after the interruption toward the task relevant movie, and whether continuity and color conditions before the interruption had an influence on the reorientation of attention toward the task-relevant movie. Two main dependent variables are discussed: the first main dependent variable is the percentage of trials where the first saccade is directed toward the task-relevant video, and the second main dependent variable is the SRTs that show the time required for the subjects to make the first saccade toward the task-relevant video, immediately when the videos reappeared after the interruption, respectively after the cut. Whenever the assumption of sphericity is violated (indicated by Mauchly test), the Greenhouse-Geisser correction of sphericity for degrees of freedom is applied.

5.2 Dependent variables

After the significant oculomotor events such as saccades, fixations and blinks are defined for every trial, the viewer's first saccades on the task-relevant or the task-irrelevant video were determined. Saccades that fell inside invisible areas of interest (AOIs) were detected, with modest extension outside the borders of both videos. The starting point is considered at landing and fixation of the first saccade toward the task-relevant video, after each cut. The percentages of trials in which the first saccade landed in the task-relevant video were analyzed.

A three way repeated measurement ANOVA was conducted to compare the means of error rates between the main conditions of continuity, color and location. Every main effect for error rates was separately elaborated.

5.2.1 First saccades toward task-relevant video

5.2.1.1 The main effect of continuity

Statistical analysis show significant main effect for continuity condition with $F(2, 70) = 44.116$, $p = .000$.

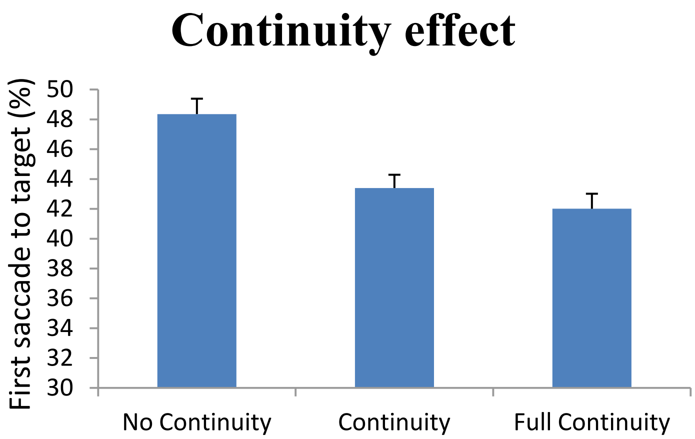

Figure 5.Percentage of first saccades to no continuity, continuity and full continuity conditions.

Figure 5 shows that the first saccades were fastest when the scene reappeared as full continuous after the cut occurred *(M = 42.00, SD = 1.00)*. On the other side, the slowest saccades were performed when the scene reappeared as no continuous after the cut *(M = 48.35, SD = 1.03)*. The intermediate saccades were performed during the continuous trials *(M = 43.39, SD = .89)*. Error bars represent the 95% confidence intervals.

Pairwise comparisons were conducted between the conditions of no continuity, continuity and full continuity, to see whether there were any statistically significant effects between these conditions. Based on the values, the no continuity is statistically significant with continuity and full continuity. However, the continuity is not statistically significant with full continuity. This means that the changes are likely to occur due to chance and not due to manipulation.

5.2.1.2 The main effect of color

A statistically significant main effect of color condition was found with *F (3,105) = 15.523, p = .000.*

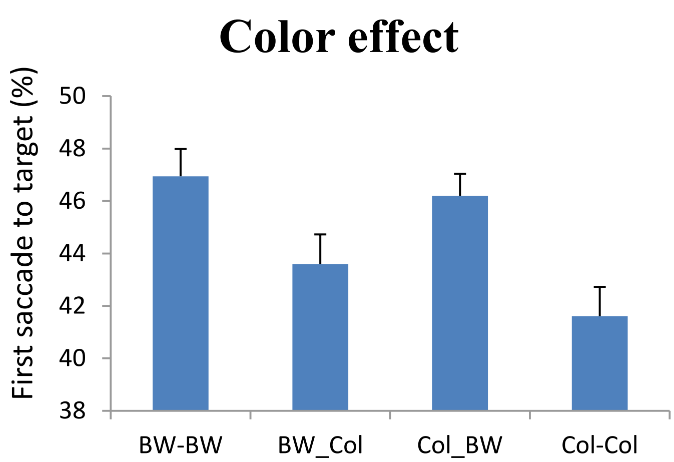

Figure 6. Percentage of first saccades among black and white to black and white, black and white to color, color to black and white and color to color conditions.

Figure 6 shows that the first saccades in color error rates were fastest when the scene reappeared from color to color after the cut occurred *(M = 41.60, SD =1.12)*. However, the slowest saccades are shown when the trial changes from black and white to black and white *(M = 46.94, SD = 1.04)*. Results show that Black and white to Color effect saccades after the interruption *(M=43.59, SD=1.13)* and Color to Black and white effect after the cut *(M=46.19, SD=.83)*. Error bars represent the 95% confidence intervals.

Pairwise comparisons were conducted between the four levels of color, to find out whether there was any statistically significant effect between these levels. Based on the values, the black and white condition is statistically significant with black and white to color and color to color. However, the black and white condition is not statistically significant with color to black and white condition. This change can occur due to chance and not manipulation. Black and white to color is statistically significant with color to black and white condition, but it is not significant with color to color condition. Moreover, color to black and white is statistically significant to black and white condition and color to color condition.

The statistically significant changes between different conditions mean that the changes occurred due to variable manipulation and not due to chance.

5.2.1.3 The main effect of location

A statistically significant main effect for location was found, $F(1, 35) = 104.85$, $p = .000$.

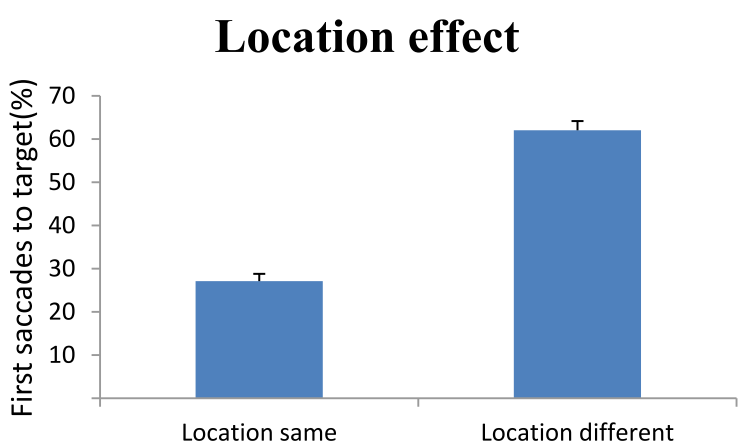

Figure 7. Percentage of the first saccades to location same and location different conditions.

Figure 7 shows that the first saccades were the fastest when scenes reappeared in the same location, before the cut occurred $(M = 27.12, SD = 1.69)$.On the other side, the slowest saccades were detected when the location changed across the trials $(M = 62.04, SD = 2.13)$. Error bars represent the 95% confidence intervals.

5.1.1.4 The main interaction between continuity and color

A statistically significant main effect between continuity and color condition was found, $F\ (6, 210) = 11.10, p = .000$.

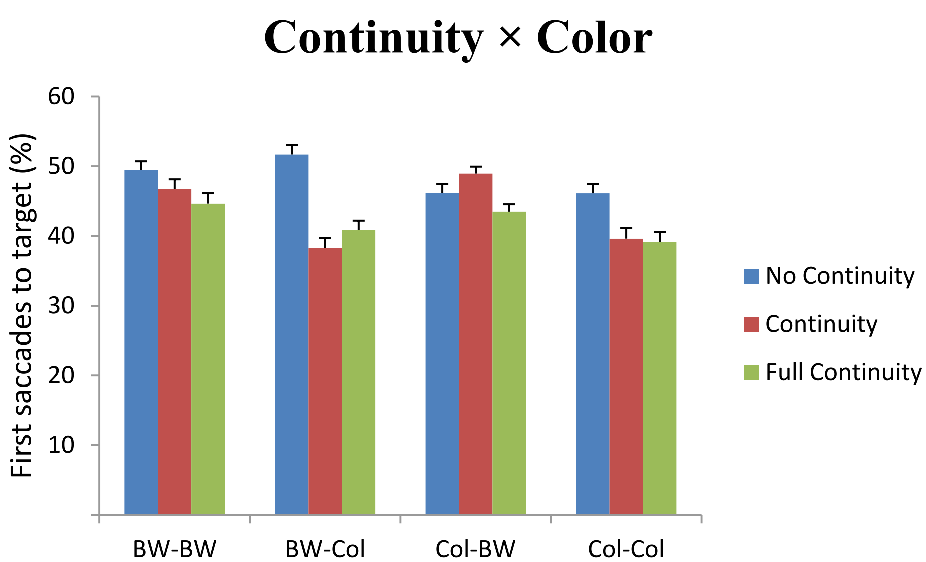

Figure 8. Percentage of the first saccades for the interaction between Continuity and Color conditions.

Figure 8 shows the percentage of the first saccades to target during the interaction between three levels of Continuity and four levels of Color in dynamic scenes. The participants performed the fastest saccades on Continuous scenes that changed from a pre-cut Black and white to Color in a post-cut scenes *(M = 38.28, SD = 1.43),* and slowest saccades when the scene changed from Black and white to Color in No Continuous scene *(M = 51.67, SD = 1.39).* Error bars represent the 95% confidence intervals. Other values are shown in table 1.

Table 1

Mean error rates during the Continuity and Color interaction

				95% Confidence Interval	
Continuity	Color	Mean	Std. Error	Lower Bound	Upper Bound
no cont	bw-bw	49.443	1.252	46.901	51.984
	bw-col	51.672	1.397	48.835	54.509
	col-bw	46.178	1.244	43.652	48.703
	col-col	46.111	1.322	43.427	48.795
cont	bw-bw	46.744	1.369	43.964	49.524
	bw-col	38.288	1.436	35.373	41.202
	col-bw	48.939	.995	46.919	50.959
	col-col	39.611	1.493	36.581	42.641
full cont	bw-bw	44.635	1.480	41.631	47.640
	bw-col	40.818	1.368	38.041	43.595
	col-bw	43.472	1.062	41.317	45.627
	col-col	39.097	1.435	36.184	42.010

5.1.1.5 The main interaction between Continuity and Location

There is no statistically significant difference between the conditions of Continuity and Location with $F(2, 70) = .010$, $p = .990$. The difference between the conditions is likely to occur due to chance, rather than manipulation.

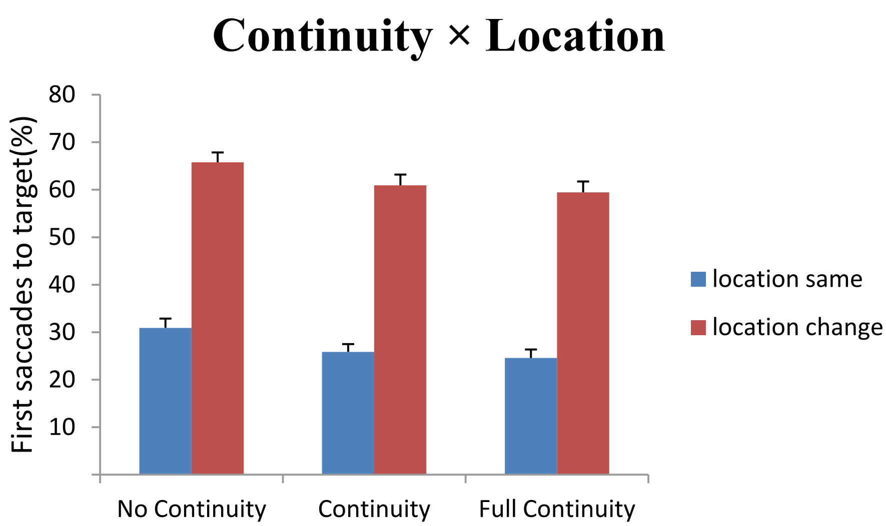

Figure 9. Percentage of the first saccades for the interaction between Continuity and Location conditions.

In the Continuity and Location interaction (fig. 9), the participants performed faster saccades when the location was the same and the Continuity repeated all features (Full Continuity) from a pre-cut to post-cut scene *(M = 24.57, SD = 1.79)*. On the other side when the Location changed from a pre-cut scene to a post-cut scene, the saccades were the fastest when it switched from Full Continuity throughout the trials *(M = 59, 44, SD = 2.29)*. The values for No Continuity and Continuity are presented in table 2.

Table 2

Mean error rates during the interaction between the Continuity and Location

				95% Confidence Interval	
Continuity	Location	Mean	Std. Error	Lower Bound	Upper Bound
no cont	same	30.919	1.954	26.953	34.886
	diff	65.782	2.064	61.592	69.972
cont	same	25.869	1.635	22.550	29.189
	diff	60.921	2.266	56.321	65.521
full cont	same	24.571	1.799	20.918	28.223
	diff	59.441	2.292	54.788	64.093

5.1.1.6 The main interaction between Color and Location

There was a main statistically significant effect for interaction between Color and Location conditions with *F (3,105) = 2.93, p = .037.*

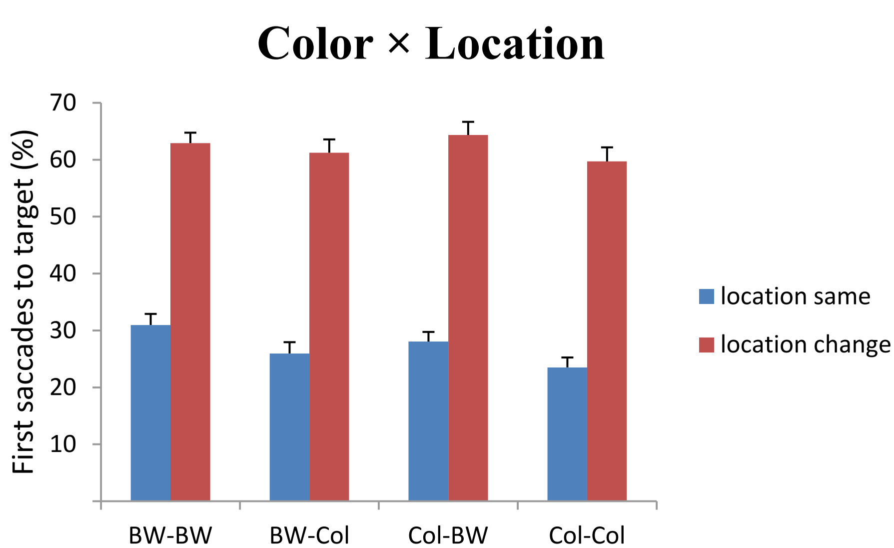

Figure 10.Percentage of the first saccades for the interaction between Color and Location conditions.

In the Color and Location interaction (fig. 10), the participants performed faster first saccades when the Location was the same and the Color repeated from a pre-cut to post-cut scene *(M = 23.51, SD = 1.74)*. On the other side when the Location changed from a pre-cut scene to post-cut scene, the first saccades were the fastest when it switched from Color to Color throughout the trials *(M = 59,69 SD = 2.47)*. Error bars represent the 95% confidence intervals. Other values are presented in the table 3.

Table 3

Mean error rates between the Color and Location interaction

				95% Confidence Interval	
Color	Location	Mean	Std. Error	Lower Bound	Upper Bound
bw-bw	same	30.963	1.947	27.010	34.916
	diff	62.918	1.835	59.193	66.644
bw-col	same	25.958	1.987	21.924	29.993
	diff	61.227	2.345	56.466	65.988
col-bw	same	28.045	1.701	24.592	31.498
	diff	64.348	2.312	59.654	69.041
col-col	same	23.514	1.743	19.974	27.053
	diff	59.699	2.478	54.669	64.730

5.1.1.7 The main interaction between Continuity, Color and Location same

There was a main significant effect between the interaction of three conditions (Continuity, Color, same Location) and (Continuity, Color, different Location), $F(6, 210) = 7.03, p = .000.$

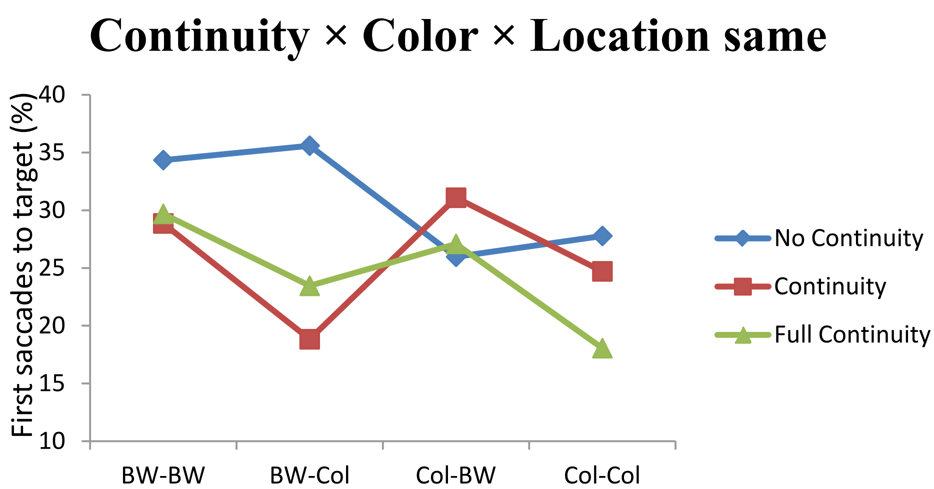

Figure 11. Percentage of the first saccades during the interaction between the conditions of Continuity, Color and same Location.

In Figure 11 the main effect of interactions between the condition of Continuity with three levels, Color with four levels and Location with one level are shown. The first saccades of the participants were fastest when the scene reappeared Full Continuous from Color to Color during the same Location *(M = 18.05, SD = 1.98)*. In contrary the slowest first saccades were performed when the scene reappeared as black and white to Color, No Continuous at the same Location *(M = 35.59, SD = 2.76)*. Other values are shown in table 4.

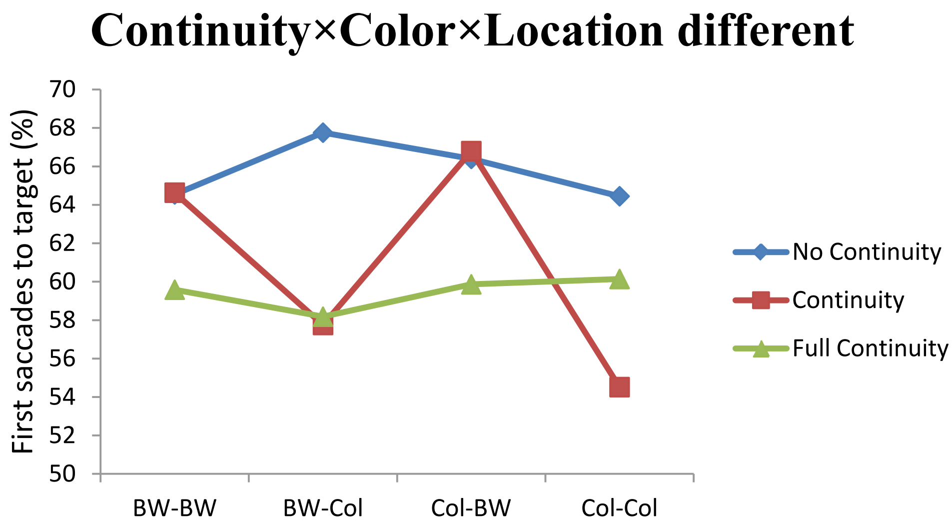

Figure 12. Percentage of the first saccade during the interaction between the conditions of Continuity, Color and different Location.

In figure 12 the main interaction between three conditions with different levels is shown. The Location is different. The participants performed the first saccade fastest when the post-cut scene reappeared as Continuous, Color to Color in different Location *(M = 54.51, SD = 3.31)*. The slowest first saccades were performed when the scene reappeared as No Continuous, Black and white to Color in different location *(M = 67.75, SD = 2.24)*.

Tabele 4

Mean error rates between the Continuity, Color and Location interaction

					95% Confidence Interval	
Continuity	Color	Location	Mean	Std. Error	Lower Bound	Upper Bound
no cont		same	34.343	1.851	30.586	38.101
	bw-bw	diff	64.542	2.326	59.821	69.264
		same	35.590	2.760	29.988	41.193
	bw-col	diff	67.754	2.248	63.191	72.317
		same	25.966	2.146	21.610	30.323
	col-bw	diff	66.389	2.571	61.169	71.609
		same	27.778	2.267	23.176	32.380
	col-col	diff	64.444	2.321	59.732	69.157
cont		same	28.858	2.361	24.064	33.652
	bw-bw	diff	64.630	2.122	60.322	68.937
		same	18.827	1.895	14.981	22.674
	bw-col	diff	57.749	2.833	51.998	63.499
		same	31.085	2.175	26.668	35.501
	col-bw	diff	66.793	2.103	62.524	71.062
		same	24.708	1.976	20.696	28.719
	col-col	diff	54.514	3.319	47.777	61.251
full cont		same	29.688	2.556	24.498	34.877
	bw-bw	diff	59.583	2.257	55.001	64.165
		same	23.457	2.226	18.938	27.975
	bw-col	diff	58.179	2.567	52.968	63.390
		same	27.083	2.038	22.946	31.221
	col-bw	diff	59.861	2.698	54.383	65.339
		same	18.056	1.989	14.019	22.093
	col-col	diff	60.139	2.531	55.000	65.278

5.2 Analysis of the saccadic reaction times (SRTs)

For the analysis of SRTs there were in total 37 trials (0.2%) excluded, in which there was no saccade detected to the task-relevant movie. The remaining data was considered valid when the SRTs in a trial differed by no more than ±2 SD around the mean of SRTs in the subject of the respective condition. A total of 800 trials (4.8% of the full data set) were excluded due to this criterion. The remaining data was aggregated again over participants and mean values were submitted to repeated measures ANOVAs. In a supplementary analysis of SRTs only trials when the target movie was fixated with the first saccade were analyzed. For that purpose when the first saccade went to the distractor movie instead of the target movie, another 6,649 trials (40.5%) were excluded. Similar procedure for ANOVA results was applied as for the first saccades toward the task-relevant video.

The results of ANOVA table (Within- subjects Effects) with corrected F-values, shows that the significant value in all conditions is less than .05, so we can conclude that there is a statistically significant difference between three main conditions and their interaction. Because of this we can conclude that between some conditions applied in this experiment, there is a statistically significant effect. To find out which conditions had a statistically significant effect, we conducted a Bonferroni post hoc test. Each main effect and interaction between effects will be examined separately.

5.2.1 The main effect of Continuity

A three way repeated measures ANOVA was conducted to compare the means of saccade latency related to Continuity, Color and Location effect. A statistically significant main effect of Continuity was found with $F(2, 70) = 92.943, p = .000.$

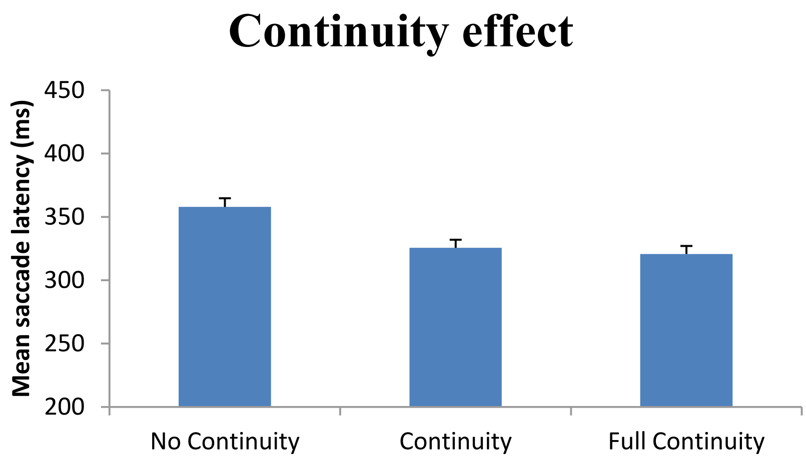

Figure 11.Mean saccade latencies among No Continuity, Continuity and Full Continuity conditions.

Figure 11 shows that with increased Continuity, the mean of saccade latency decreased. This main effect of Continuity reflects that the participants were faster to saccade to the scenes that were more continuous. They performed fastest saccades when the scene after the cut appeared as Full Continuous *(M = 320.74, SD = 6.31)*, rather than in Continuous *(M = 325.56, SD = 6.41)* and No Continuous scenes *(M = 357.85, SD = 6.85).* Error bars represent the 95% confidence intervals.

Pairwise comparisons for Continuity condition between No Continuity, Continuity and Full Continuity are compared to see whether there is a statistically significant effect between these conditions. Based on the values we can conclude that SRTs in the No Continuity conditions is statistically slower than in the other two conditions. This means that the difference between conditions of No Continuity is significantly different from Continuity and Full Continuity in terms of slower SRTs performed after the cut occurred. The difference is probably due to the effect of manipulation and not due to chance. However, the Continuity condition is not statistically different in SRTs compared to Full Continuity conditions; therefore the difference is likely to be due to chance, and not due to manipulation of conditions.

5.2.2 The main effect of Color

There was a significant main effect of Color with F (3,105) = 19.074, p = .000 in a within subject effect table.

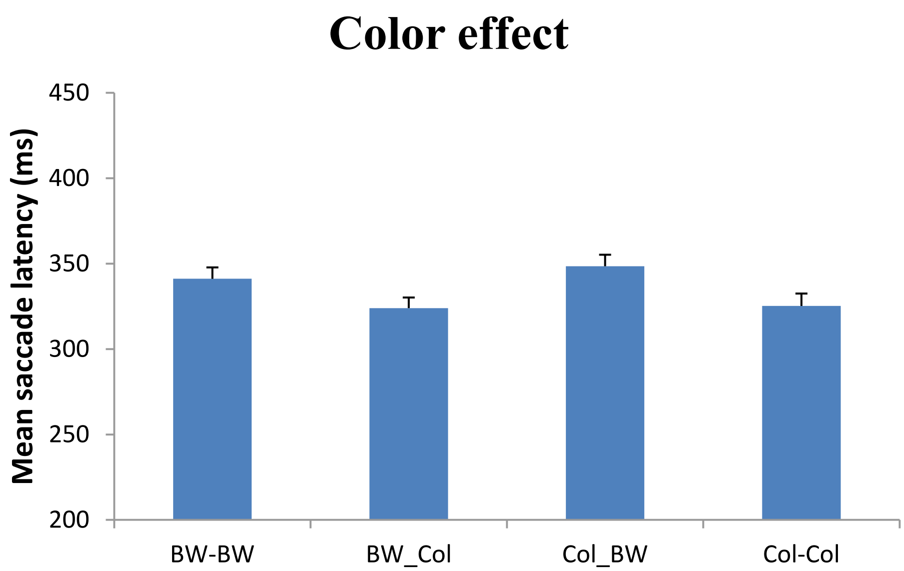

Figure 12. Mean saccade latencies among black and white to black and white, black and white to color, color to black and white and color to color conditions.

Figure 12 shows that the participants performed fastest in trials when the black and white scenes reappeared in color after the cut *(M = 323.98, SD = 6.22)*. The slowest saccades were in the trials when the color pre-cut scene appeared as black and white after the cut *(M = 348.49, SD = 6.75)*. Other values were in scenes from black and white to black and white *(M = 341.14, SD = 6.67)*, and color to color *(M = 325.26, SD = 7.26)*. Error bars represent the 95% confidence intervals.

Pairwise comparisons for color effect between black and white to black and white, black and white to color, color to black and white and color to color effects were compared to find out whether there is a statistically significant difference between them. The results show that black and white to black and white is statistically different on faster SRTs compared with black and white to color and color to color effects, but it did not show any statistically significant difference compared with color to black and white effect. On the other side, black and white to color effect show statistical difference on faster SRTs compared with color to black and white, but did not show any statistical difference with color to color effect. However, color to black and white showed statistically significant faster SRTs when compared with color to color effect.

5.2.3 The main effect of location

A significant main effect of location was found in within subject effects table with *F (1, 35) = 99,364, p=.000.*

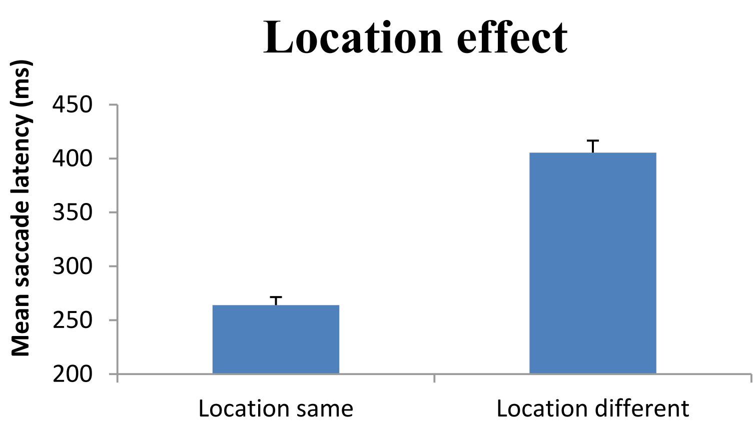

Figure 13. Mean saccade latencies during the location same and location different conditions.

The location figure 13 shows that the participants performed faster SRTs when the scene appeared on the same side after the cut *(M = 263.98, SD = 7.47)* and a slower SRTs when the scene appeared on a different side *(M = 405.45, SD = 11.14).* Error bars represent the 95% confidence intervals.

Pairwise comparisons for location conditions between the location same and location different are compared to see whether there is a statistically significant difference. The values show that location same is statistically different on performing faster SRTs compared with location different. For this reason, we can conclude that the location same and location different are statistically significant in terms of faster SRTs performed after the cuts.

5.2.4 The main interaction between Color and Continuity

Statistical analysis yielded significantly main effect between the Continuity and Color conditions with $F(6,210) = 8.070, p = .000$.

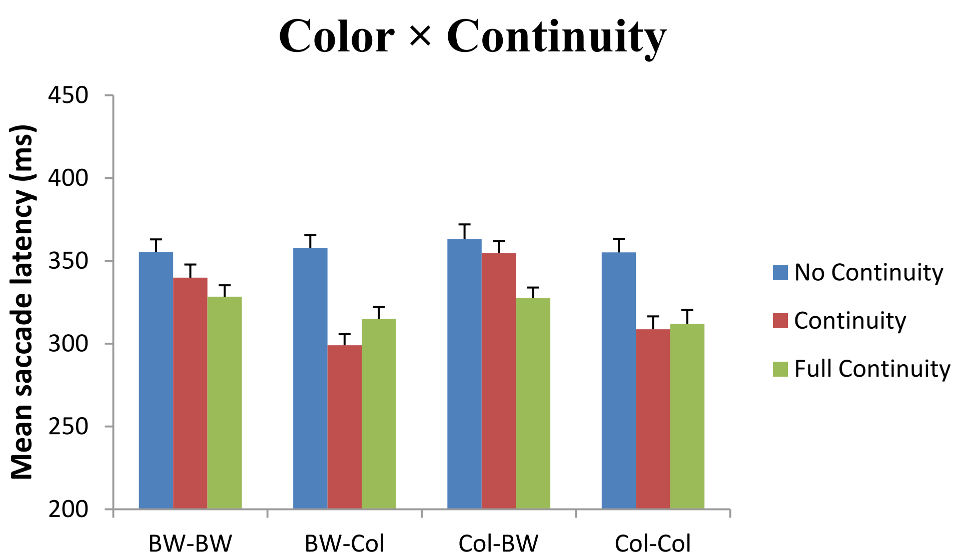

Figure 14. Mean saccade latencies during the interaction between Continuity and Color conditions.

Figure 14 shows the mean saccadic latencies during the interaction between three levels of Continuity and four levels of Color in dynamic scenes. The participants performed the fastest SRTs on Continuous scenes that changed from pre-cut black and white scene to Color in the post-cut scene *(M = 299.05, SD = 6.68)*, and slowest SRTs when the scene changed from Color in No Continuous scene to black and white in No Continuous scene *(M = 363.22, SD = 8.82)*. Error bars represent the 95% confidence intervals. Other values are shown in table 6.

Table 6

Continuity and Color effect

Continuity	Color	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
no cont	bw-bw	355.249	7.752	339.511	370.987
	bw-col	357.825	7.735	342.123	373.527
	col-bw	363.227	8.822	345.318	381.135
	col-col	355.110	8.264	338.333	371.887
cont	bw-bw	339.862	7.967	323.688	356.037
	bw-col	299.057	6.682	285.493	312.621
	col-bw	354.631	7.343	339.724	369.537
	col-col	308.716	7.796	292.888	324.544
full cont	bw-bw	328.336	6.953	314.221	342.452
	bw-col	315.059	7.251	300.340	329.779
	col-bw	327.614	6.317	314.789	340.439
	col-col	311.968	8.527	294.658	329.279

5.2.5 The interaction between Continuity and Location

There was a statistically significant interaction effect between the Continuity and Location condition with $F(2, 70) = 7.939, p = .001$.

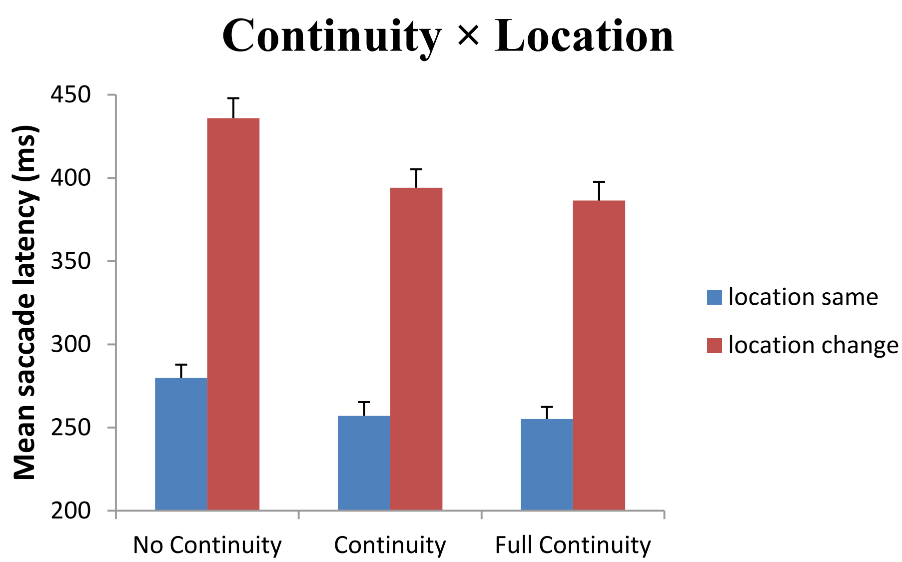

Figure 15.Mean saccade latencies during the interaction between the Continuity and Location conditions.

Between the Continuity and Location interaction (fig. 15), the participants performed faster SRTs in three levels of Continuity when the scene reappeared in the same Location. Moreover participants showed the fastest SRTs when the post-cut scene repeated all features from the pre-cut scene (Full Continuity) to post-cut scene in the same Location *(M = 255.12, SD = 7.26)*. Error bars represent the 95% confidence intervals. Other values are presented in the table 7.

Table 7

Continuity and Location effect

				95% Confidence Interval	
Continuity	Location	Mean	Std. Error	Lower Bound	Upper Bound
no cont	same	279.798	8.018	263.521	296.076
	diff	435.907	11.988	411.571	460.243
cont	same	257.028	8.249	240.282	273.775
	diff	394.105	11.088	371.595	416.614
full cont	same	255.126	7.261	240.386	269.866
	diff	386.363	11.296	363.431	409295

.

5.2.6 The interaction between Color and Location effect

There was a statistically significant interaction effect between Color and Location conditions with *F (3,105) = 4.442, p = .006*

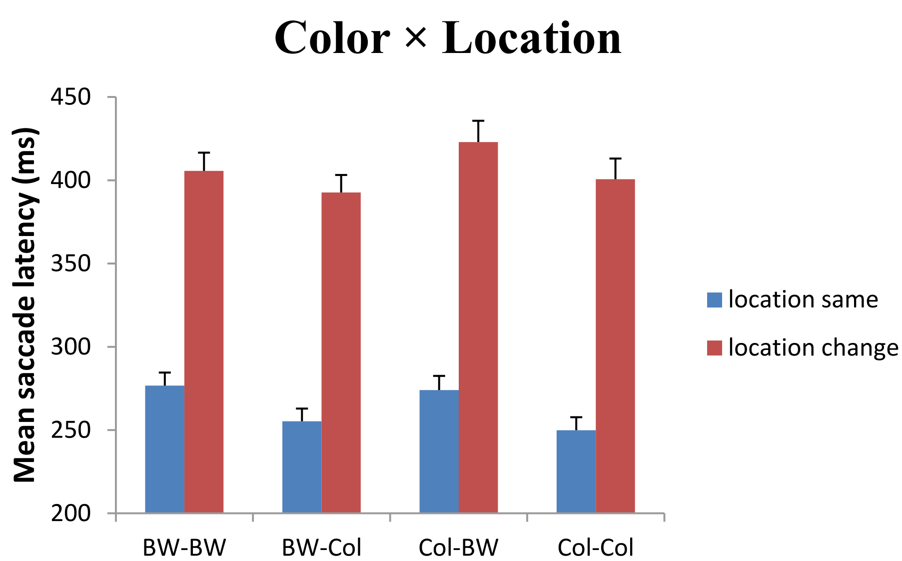

Figure 16. Mean saccade latencies during the interaction between Color and Location conditions.

In the Color and Location interaction (fig. 16), the participants performed faster SRTs when the Location was the same and the Color repeated from a pre-cut to post-cut scene *(M = 249.93, SD = 7.80)*. On the other side when the Location changed from a pre-cut scene to post-cut scene, the SRTs were the fastest when it switched from Black and white to Color throughout the trials *(M = 392.67, SD = 10.50).* Error bars represent the 95% confidence intervals. Other values are presented in the table 8.

Table 8

Color and Location effect

Color	Location	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
bw-bw	same	276.693	7.854	260.749	292.638
	diff	405.605	10.969	383.336	427.874
bw-col	same	255.286	7.639	239.778	270.794
	diff	392.675	10.500	371.359	413.991
col-bw	same	274.027	8.513	256.745	291.309
	diff	422.954	12.759	397.052	448.855
col-col	same	249.930	7.809	234.077	265.784
	diff	400.599	12.445	375.336	425.863

5.2.7 The interaction between Continuity, Color and Location effect

There was a significant interaction between the three conditions (Continuity, Color, same Location) and (Continuity, Color, different Location), with $F(4.29, 150.37) = 9.032, p = .000.$

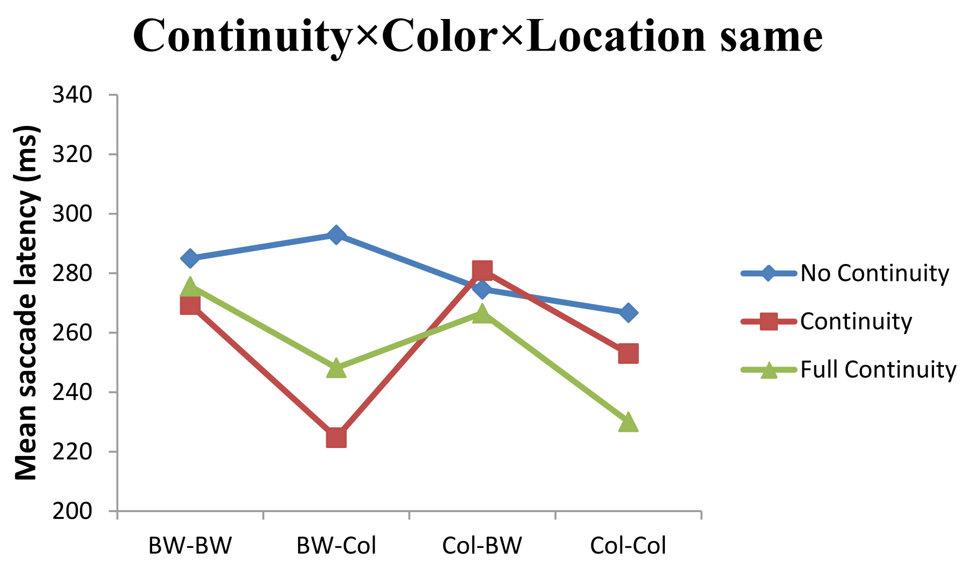

Figure 17. Mean saccade latencies during the interaction of Continuity, Color and same Location conditions.

In figure 17 the interactions between the condition of Continuity with three levels, Color with four levels and Location with one level are shown. Participants' SRTs were fastest when the scene reappears Continuous and in Black and white due to the same Location *(M = 224.72, SD = 7.97)*. In contrary the slowest SRTs are performed when the scene reappeared as Black and white, No Continuous at the same Location *(M = 292.91, SD = 9.99)*. Other values are shown in table 16.

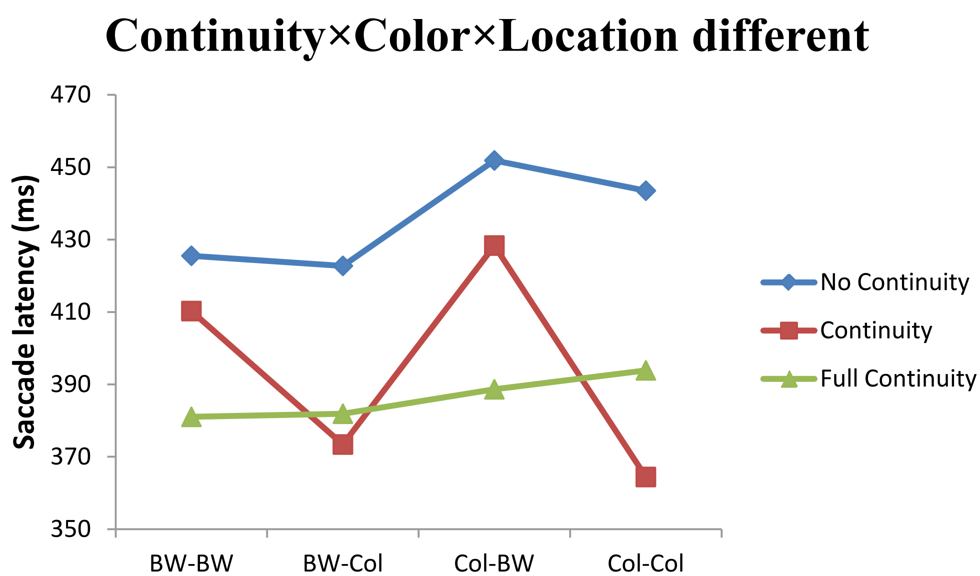

Figure 18. Mean saccade latencies during the interaction of Continuity, Color and different Location conditions.

Figure 18 shows the interaction between three conditions with different levels. The Location is different. The participants performed the fastest SRTs when the post-cut scene reappeared as Continuous, Color to Black and white in a different Location. The slowest SRTs were performed when the scene reappeared as No Continuous, Black and white to Color with a different location.

Tabele 9

Interaction between Continuity, Color and Location

					95% Confidence Interval	
Continuity	Color	Location	Mean	Std. Error	Lower Bound	Upper Bound
no cont	bw-bw	same	284.973	8.317	268.089	301.857
		diff	425.525	13.058	399.015	452.034
	bw-col	same	292.910	9.999	272.610	313.209
		diff	422.740	11.760	398.866	446.614
	col-bw	same	274.595	11.524	251.200	297.991
		diff	451.858	14.789	421.834	481.882
	col-col	same	266.715	8.446	249.569	283.861
		diff	443.505	14.502	414.064	472.946
cont	bw-bw	same	269.472	10.409	248.340	290.604
		diff	410.253	12.408	385.063	435.443
	bw-col	same	224.724	7.976	208.532	240.915
		diff	373.391	11.333	350.383	396.398
	col-bw	same	280.904	9.873	260.861	300.946
		diff	428.357	13.606	400.736	455.979
	col-col	same	253.014	9.879	232.959	273.069
		diff	364.418	12.195	339.660	389.176
full cont	bw-bw	same	275.635	8.840	257.690	293.581
		diff	381.038	10.799	359.115	402.961
	bw-col	same	248.224	8.306	231.362	265.086
		diff	381.895	11.307	358.940	404.850
	col-bw	same	266.583	9.823	246.640	286.525
		diff	388.645	12.075	364.131	413.159
	col-col	same	230.061	7.716	214.397	245.726
		diff	393.875	14.702	364.028	423.723

6. Discussion

In this Master thesis the factors that drive the eye movements while perceiving edited dynamic scenes are discussed. The effectiveness of gaze reallocation towards the task relevant movies after the interruptions is measured in an eye tracking experiment. Results of the present study show that the attention is drawn by repeated features among cuts.

This study has proved that repetition of more features from a pre-cut scene to a post-cut scene triggers a faster gaze reallocation. The major independent variable is similarity between pre and post-cut images, both in terms of rough content-wise scene or action continuation across the cut, as well as, in terms of measured color similarity between the pre and post-cut image. The major dependent variables are the error rates of first saccade toward the target movie, and the saccadic reaction times of a position switch toward a target movie, when both sport videos were shown in parallel at the same time.

Viewers effectively reallocated their gaze toward the task-relevant movie, when the visual scene after the cut reappeared as continuous (Ansorge, Buchinger, Valuch, Patrone & Scherzer, 2014). The present data reconfirm the assumption by Ansorge et al., (2014) that subjects performed fastest SRTs in the scenes when Full Continuity effects were applied after the interruption, independently whether the Color condition was present before the interruption. Furthermore, Kristjánsson et al., 2002 found that the gaze is effectively reallocated and benefits from color condition after the interruption. In our study, the color seems to play an effective role in facilitating the saccades toward discriminating the irrelevant from the relevant scene reappearing after the interruption. The present data show that viewers performed slightly faster SRTs in a Full Continuity condition when the video after the interruption was shown in Color rather than in Black and white condition. Based on results from Maljkovic and Nakayama (1994), short-term memory plays a main role on effective processing when the Color condition is repeated from the previous trial after the interruption. Nevertheless, we cannot conclude that the Color condition plays the main role in memory encoding and retrieval. The present data show that even among the Black and white to Black and white cuts, faster SRTs were performed during the Full Continuity comparing to No Continuity conditions. The third condition for effective eye reallocation following a cut is the location condition (Wolfe, 1994). Based on the present data, viewers performed faster SRTs during the Continuity and Color conditions whenever the video after the interruption was shown at the same Location.

A major finding is that visual attention is re-oriented faster to the relevant movie when more elements are repeated after the interruption. In a study by Itti and Baldi (2009) is assumed that the best predictors for attracting attention after the cut are novel features, or surprised visual stimuli. Adapting this model to our study, the assumption would be that a No Continuity scene draws more visual attention. In contrary with this inference our results are consistent with (Kristjánsson & Campana, 2010; Valuch et al., 2013; Valuch et al., 2014; Ansorge et al., 2014) studies, assuming that viewers are focused on finding repeated features from a pre-cut scene. Ansorge et al., (2014) propose a two-step model of studying attention in edited dynamic scenes. They assume that a higher reciprocal influence exists between memory and saccades. Perceiving a movie sequence requires different gaze strategies comparing with the perception of sequence reappearing after the interruption. During the perception of a movie sequence the novel stimuli are fixated and encoded into memory, while after the interruption the process of recognizing plays the main role. The subjects' task is to recognize whether the following sequence is Continues or No continues from the previous sequence. The memory representations are required in order to search for visual similarities between previous and actual sequence (feature priming). All in all, the two-step model assumes that feature priming model plays the main role in perceiving the scenes across the cuts, however, in between cuts the viewers preferred novel stimuli. In addition, Valuch et al., (2013) argue that when viewers try to recognize an image, they prefer looking at repeated visual content or at positions they were attending before. However, viewers have only a short time after the cut on disposal, so they are focused on finding the object similarities from a previous scene (Valuch et al., 2014). Our results proved that the feature priming model is more effective on attentional reallocation following a cut, rather than surprise model of attention.

Since most of the visual studies are on how humans perceive static scene images (Henderson and Hollingworth, 1999) this study gives an insight on the ecological approach or perception of visual dynamic scenes. Based on the study by Hochberg and Brooks (1996) perception is a mental representation; therefore humans perceive film scenes by creating mental representations and anticipating them afterwards. In contrary, Gibson's ecological approach assumes that humans perceive dynamic scenes directly, with a minimum of attentional effort

(Gibson, 1978). Our study is focused on supporting further research in the field of more natural scenes.

Although movies contain many cuts during the editing process, movie editors very frequently use editing techniques to encourage the Continuity in order to maintain the narrative in the scenes. Movie scenes do not have the same duration and are presented with many camera shots, starting with longer shots at the beginning, shorter in the middle and longer at the end. This rise and fall of shots guides viewers' attention (Smith & Henderson, 2008; Smith et al., 2012). Another important aspect of Continuity that is used by film editors for advancing the narrative in films on Continuity editing rules is the "180 degree rule" which ensures spatial Continuity (Bordwell & Thompson, 2001). Based on actual results, a possible explanation of how the movies are perceived might somehow be similar to what we observe in visual search experiments ("inter-trial priming"). For instance, the subject has to search for something that he/she needs to find in the first trial, then he/she finds it, and in the second trial the scene changes, so the subject has to search for something else again. The same procedure occurs in the third trial and so on. Similarly, it could be that the viewers in film want to connect two scenes with each other by finding the objects that repeat from the first scene before the cut was performed.

A previous study by Smith and Henderson (2008) found that scenes following Continuity editing rules such as within scene editing increases the edit blindness significantly, compared to decreased edit blindness in between scene editing. Similar to this finding, our results speculate that the SRTs were significantly faster in scenes repeating features after the cut (within scene cut), in comparison to slower SRTs performed during the discontinuous scenes (between scenes cut), where most of the features did not repeat to the post-cut scene. Based on current results, we can speculate that in between scenes cut more time might be required to process the visual information rather than within scenes cut.

The interdisciplinary aspect of this study concerns the different disciplines and their different methodological approaches related to the present research question. This study takes into consideration on the one hand attention as a phenomenon studied by cognitive psychology, and on the other hand relates it to dynamic scene perception such as a film theory phenomenon.

In most cases film editors cut the scenes intuitively without taking into consideration the exact times at which the scene cut should be applied (Smith and Henderson, 2008). By taking into consideration the variables that attract attention after the cuts in dynamic scenes and facilitate Continuity during cuts, this study provides more insight for film editors.

This study will contribute to researching systematically the factors that contribute to film perception. The insights of this study thus cannot only help to better understand the mechanisms of visual attention in more ecological viewing situations (such as film viewing), but also inform cognitive film theorists in various other applications, such as developments in the field of human computer interaction (video gaming, the browsing of web pages etc.). Limits of our experimental design in comparison to more ecological viewing situations is because our experimental design enables precise measurements of the viewer's first eye movements and saccadic latency following the cuts, but is different from ecological viewing situations in which an uninterrupted complete film is perceived.

For future research it would be interesting to investigate whether there would be different results, when participants watch full screen movies, rather than watching target and distractor movies shown simultaneously. Displaying only one movie in a full screen display can show different results in comparison with the current study, where both the target and distractor movies were presented at the same time. Moreover, showing a complete movie would be interesting as well, rather than showing short scenes with many cuts. A further interesting investigation would be to show films with a soundtrack rather than silent scenes. Presenting films with a soundtrack might provide different results compared to silent scenes as presented in current study. Finally, the duration of memory representation while watching an edited movie remained unexplored in this study. These questions remain to be clarified by future research.

7. Conclusio

The present study investigated how the attention is reoriented after the interruption toward the target movie while manipulating the conditions of Continuity, Color and Location from pre-cut to post-cut visual scenes. The results confirm that the reorientation of viewers' attention was facilitated after the interruption, when most of the features are repeated from a pre-cut to a post cut scene. Memory seems to be one of the factors that play crucial role during the perception of continuity editing scenes. Moreover, viewers are attracted more by bottom up features that are repeated from a pre-cut to post-cut scene.

The present results pave the way to future research not only on film perception, but also contribute to better understand the human computer interaction.

References

Anderson, J. D. (1996). *The reality of Illusion: An Ecological Approach to Cognitive Film Theory.* Carbondale, USA: Southern Illinois University Press

Ansorge, U., Buchinger, S., Valuch, C., Patrone, A. R., & Scherzer, O. (2014). Visual attention in edited dynamical images. *Proceedings of the 11th International Conference on Signal Processing and Multimedia Applications*, 198–205. doi:10.5220/0005101901980205

Becker, S. (2008a). Can intertrial effects of features and dimensions be explained by a single theory? *Journal of Experimental Psychology. Human Perception and Performance*, *34*(6), 1417-1440. doi:10.1037/a0011386

Becker, S. (2008b). The stage of priming: Are intertrial repetition effects attentional or decisional? *Vision Research*, 48(5), 664-684. doi: 10.1037/a0011386

Becker, B., & Cassimeris, L. (2005). Cytoskeleton: microtubules born on the run. *Current Biology□: CB*, *15*(14), R551–4. doi:10.1016/j.cub.2005.07.006

Becker, S. & Horstmann, G. (2009). A feature-weighting account of priming in conjunction search. Attention, *Perception, & Psychophysics*, 71 (2), 258-272. doi: 10.3758/APP.71.2.258

Becker, S., Valuch, C., & Ansorge, U. (2014). Color priming in pop-out search depends on the relative color of the target. *Frontiers in Psychology*, Volume 5. doi:10.3389/fpsyg.2014.00289

Bordwell, D., & Thompson, K. (2001) *Film Art: An introduction:* The relation of shot to shot: Editing (6th ed.,). New York, USA: Mc Graw Hill.

Born, S., Ansorge, U., & Kerzel, D. (2012). Feature-based effects in the coupling between attention and saccades. *Journal of Vision*, *12*, 1–17. doi:10.1167/12.11.27.

Cohen, M., Cavanagh, P., Chun, M., & Nakayama, K. (2012). The attentional requirements of consciousness. *Trends in Cognitive Sciences*, *16*(8), 411–417. doi:10.1016/j.tics.2012.06.013

Cornelissen, F., Peters, E., & Palmer, J. (2002). The Eyelink Toolbox: eye tracking with MATLAB and the Psychophysics Toolbox. Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society, Inc, 34(4), 613–617. doi:10.3758/BF03195489

Cutting, J. E. (2005). Perceiving scenes in film and in the world. J. D. Anderson & B. F. Anderson (Eds.) *Moving image theory: Ecological considerations* (pp. 9-27). Carbondale, IL: University of Southern Illinois Press.

Cutting, J. E., Brunick, K. L., Delong, J. E., Iricinschi, C., & Candan, A. (2011). Quicker, faster, darker: Changes in Hollywood film over 75 years. *I-Perception,* 2(6), 569–576. doi:10.1068/i0441aap

Cutting, J. E., DeLong, J. E., & Nothelfer, C. E. (2010). Attention and the evolution of Hollywood film. *Psychological Science*, *21*(3), 432–9. doi:10.1177/0956797610361679

Deubel, H., & Schneider, W. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*(12), 1827-1837. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/0042698995002944

Findlay J.M. (2005). Covert attention and saccadic eye movements. In Itti, L., Rees, G. & Tsotsos, J.(Eds.), Neurobiology of Attention (114-117). London New York: Elsevier Academic Press.

Frank, M., Vul, E., & Johnson, S. (2010). Development of infants' attention to faces during the first year. *NIH Public Access,* 110(2), 160–170. doi:10.1016/j.cognition.2008.11.010.

Gegenfurtner, K. R., & Rieger, J. (2000). Sensory and cognitive contributions of color to the recognition of natural scenes. *Current Biology,* 10(13), 805–808. doi:10.1016/S0960-9822(00)00563-7

Gibson, J. (1978). The Ecological Approach to the Visual Perception of Pictures. *Leonardo*, *11*, 227–235. doi:10.2307/1574154

Goldstein, A. G., Chance, J. E., Hoisington, M., & Buescher, K. (1982). Recognition memory for pictures: Dynamic vs. static stimuli. *Bulletin of the Psychonomic Society*, *20*(1), 37–40. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1983-11900-001&site=ehost-live

Henderson, J. M., & Hollingworth, a. (1999). The Role of Fixation Position in Detecting Scene Changes Across Saccades. *Psychological Science*, *10*(5), 438–443. doi:10.1111/1467-9280.00183

Hillstrom, a P., & Yantis, S. (1994). Visual motion and attentional capture. *Perception & Psychophysics*, *55*(4), 399–411. doi: 10.3758/Bf03205298

Hoffman, J. (1975). Hierarchical stages in the processing of visual information. Perception & Psychophysics, 18(5), 348–354. doi:10.3758/BF03211211

Hochberg, J., & Brooks, V. (1996). *The Perception of Motion Pictures.* In E.C.Carterette & M.P. Friedman (Eds.), Handbook of Perception (205-278). New York: Academic Press.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance,* 28 (1), 113–136. doi:10.1037/0096-1523.28.1.113

Hunt, A. R., & Kingstone, A. (2003). Covert and overt voluntary attention: linked or independent? *Cognitive Brain Research*, *18*(1), 102–105. doi:10.1016/j.cogbrainres.2003.08.006

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306. doi:10.1016/j.visres.2008.09.007

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10-12), 1489–1506. doi:10.1016/S0042-6989(99)00163-7

Itti, L., Koch, C,. & Niebur, E. (1998). A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. IEEE transactions on pattern analysis and machine intelligence, 20(11), 1254-1259

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review*, *111*(4), 880–913. doi:10.1037/0033-295X.111.4.880

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition,* 22(6), 657–672. doi:10.3758/BF03209251

Mital, P. K., Smith, T. J., Hill, R. L., & Henderson, J. M. (2010). Clustering of Gaze During Dynamic Scene Viewing is Predicted by Motion. *Cognitive Computation*, *3*(1), 5–24. doi:10.1007/s12559-010-9074-z

Nothdurft, H. C., Pigarev I.N., & Kastner S. (2009). Overt and covert visual search in primates: Reaction times and gaze shift strategies. *Journal of Integrative Neuroscience,* 8(2), 137-174.

Klein, R.M. (1980). *Does Oculomotor Readiness Mediate Cognitive Control of Visual Attention?* In Nickerson, R.(Eds.), Attention and Performance VIII (259–276). Lawrence Erlbaum Associates, Inc.

Klein, R.M. & Pontefract, A. (1994). *Does oculomotor readiness mediate cognitive control of visual attention?* Revisited! In Umilta C., & Moscovitch, M. (Eds.), Attention and Performance. Consious and Nonsconscious Informaiton Processing (333- 350). Cambridge: The International Association for the Study of Attention and Performance.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, *35*(13), 1897–1916. doi:10.1016/0042-6989(94)00279-U

Kristjánsson, Á., & Campana, G., (2010). Tutorial Review Where perception meets memory□: A review of repetition priming in visual search tasks. *Attention, Perception, & Psychophysics*, *72*(1), 5–18. doi:10.3758/APP.72.1.5

Kristjánsson, Á., Wang, D., & Nakayama, K. (2010). The role of priming in conjunction search. *Journal of Vision*, *1*(3), 114–114. doi:10.1167/1.3.114

Plantinga, C. (2002). Cognitive Film Theory: An Insider's Appraisal. *Cinémas: Revue D'études Cinématographiques,* 12(2), 15. doi:10.7202/024878ar

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, *32*(1), 3–25. doi:10.1080/00335558008248231

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: evidence in favor of a premotor theory of attention. *Neuropsychologia*, *25*, 31–40. doi:10.1016/0028-3932(87)90041-8

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences*, *24*(2), 113–121. doi:10.1016/S0166-2236(00)01685-4.

Smith, T. J. (2005). *An Attentional Theory of Continuity Editing* (Doctoral dissertation). Retrieved from http://hdl.handle.net/1842/1076

Smith, T. J. (2013). Watching you watch movies: Using eye tracking to inform cognitive film theory. In A.P. Shimamura (Ed.) *Psychocinematic: Exploring Cognition at the Movies* (pp. 165-191). New York, NY: Oxford University Press.

Smith, T. J., & Henderson, J. M. (2008). Edit blindness: The relationship between attention and global change blindness in dynamic scenes. *Journal of Eye Movement Research*, 2(2), 1-17.

Smith, T. J., & Henderson, J. M. (2008). Edit blindness: The relationship between attention and global change blindness in dynamic scenes. *Journal of Eye Movement Research*, *2*(2), 1-17.

Smith, T. J., Levin, D., & Cutting, J. E. (2012). A Window on Reality: Perceiving Edited Moving Images. *Current Directions in Psychological Science*, *21*(2), 107–113. doi:10.1177/0963721412437407

Solso, R.L., Maclin, O.H., & Maclin, M.K. (2008). *Cognitive psychology*. Boston, USA.: Pearson education, Inc.

Schwan, S., & Ildirar, S. (2010). Watching film for the first time: how adult viewers interpret perceptual discontinuities in film. *Psychological Science□: A Journal of the American Psychological Society / APS*, *21*(7), 970–976. doi:10.1177/0956797610372632

Shimamura, A. P., Cohn-Sheehy, B. I., & Shimamura, T. a. (2014). Perceiving movement across film edits: A psychocinematic analysis. *Psychology of Aesthetics, Creativity, and the Arts*, *8*(1), 77–80. doi:10.1037/a0034595

Theeuwes, J. (1993). Visual selective attention: A theoretical analysis. *Acta Psychologica*, *83*(2), 93–154. doi:10.1016/0001-6918(93)90042-P

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77–99. doi:10.1016/j.actpsy.2010.02.006

Treisman, A., & Gelade, G. (1980). A Feature-Integration Theory of Attention. *Cognitive Psychology,* 12, 97–136.

Columbia, B., & Nagle, M. (1980). A Feature-Integration of Attention, *136*, 97–136.

Valuch, C., Ansorge, U., Buchinger, S., Patrone, A., & Scherzer, O. (2014). The effect of cinematic cuts on human attention. *Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video - TVX '14*, 119–122. doi:10.1145/2602299.2602307

Valuch, C., Becker, S. I., & Ansorge, U. (2013). Priming of fixations during recognition of natural scenes. *13(3):3, 1–22*. doi:10.1167/13.3.3

Wolfe, J. M. (1994). *Guided Search 2.0 A revised model of visual search.* Psychonomic Bulletin & Review, 1(2), 202-238

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: On the contributions of top-down and bottom-up guidance in visual search for feature singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(2), 483-502. doi:10.1037/0096-1523.29.2.483

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*, 1–7. doi:doi:10.1038/nrn1411

Curriculum Vitae

Personal information

Name	Blerim Zeqiri
Date of birth	25.08.1982
Place of birth	Prishtina, Kosovo

Education

since 10.2010	Middle European Interdisciplinary Master Programme in Cognitive Science, University of Vienna and Medical University of Vienna, Austria
2002 – 2005	Bachelor of Psychology, University of Pristina, Kosovo

Research experience

10.2011 - 01.2012	Conducting of Eye Tracking experiment (Suppression of memories via Think /No Think paradigm) which explored human capability of suppressing unwanted memories during the face perception task (Supervisor: Dr. Mihaly Racsmany, Budapest University of Technology and Economics, Department of Cognitive Science, Budapest, Hungary).
01.03 – 31.05. 2013	Conducting of Social neuroscience experiment, evaluating gender differences regarding the effects of stress on empathy and emotional self-other distinction abilities. (Supervisors: Mag. Livia Tomova and MSc. Irene Stepniczka, University of Vienna, Faculty of Psychology, Vienna, Austria).
01.2013 - 04.2013	Conducting the Master thesis experiment on "Priming of visual attention in dynamic visual scenes" (Supervisor: Univ. Prof. Dr. Ulrich Ansorge, University of Vienna, Faculty of Psychology, Vienna, Austria)

Languages

Albanian Language (Native)

English Language (Advanced level)

German Language (Advanced level)

Serbian Language (Advanced level)