Aesthetics of Stereoscopic Cinema

Barbara Flueckiger

Abstract: Although stereoscopic cinema was invented very early in the history of film, it did not become the standard for cinematic representations. With the latest digital wave of stereoscopic 3D cinema many shortcomings of earlier technologies have been eliminated, but debate remains about the aesthetic principles of stereoscopy. This article explores and evaluates basic approaches to aesthetic design in stereoscopic films.

Keywords: deep focus, depth of field, negative parallax, positive parallax, stereoscopy, window violation

In 2009 Kristin Thompson asked provocatively: “Has 3D already failed?” It has not failed yet. But there remain many open questions regarding the aesthetics and conventions of stereoscopic films (S3D). Discussing these questions not only sheds light on the possibilities and limitations of S3D but also allows us, on a higher level, to think about the differences between the perception of objects and images.

In his seminal study *The Photoplay* Hugo Münsterberg investigated this relationship and the differences between natural vision and the perception of moving images. On the perception of depth and movement in the movies he remarked: “Depth and movement alike come to us in the moving picture world, not as hard facts but as a mixture of fact and symbol. They are present and yet they are not in the things. We invest the impressions with them” (Münsterberg 1916: 71). He proposed a constructivist approach to the viewer’s perception and the cognitive activities that assemble meaning based on the representations displayed by the film. “Fact,” in Münsterberg’s view, refers to hard-wired forms of perception where the physiological foundations of the visual system define to a high degree what we are seeing. Conversely, “symbol” refers to learned reactions, whether they are acquired in our everyday interaction with our environment or whether they are shaped by the culture we experience in our society. Münsterberg thus addressed the crucial framework for the investigation of the aesthetics of stereoscopic cinema, namely the tension between natural perception and the perception of film as an art form strongly informed by established codes and conventions.

doi: 10.3167/proj.2012.060104 ISSN 1934-9688 (Print), ISSN 1934-9696 (Online)

The concept of aesthetics referred to in this article is a phenomenological one. Therefore the aesthetic features of a film are those that address the senses, regardless of whether they were intentionally arranged by an artistic agency, often symbolically attributed by critics to a filmmaker, or whether they are highly informed by technological possibilities and limitations. Film technology has been conceived to address the senses by arranging the stimuli in corresponding dimensions and internal formations that reflect the insights of psychophysics mainly gathered in the nineteenth and early twentieth century. With Roger Odin's semio-pragmatic approach I assume that a spectator who perceives a movie is placed in an aesthetic mode of reception by the cultural framing in which a film is presented with regard to both tradition and cultural practice. Therefore my analysis of aesthetics is located at the intersection of physiological and psychological foundations in connection with technological inventions that reflect these foundations.

Based on this framework it is the intention of this article not to establish a technological history of stereoscopic cinema, but to analyze a current practice in the newly formed digital version of it. It is the sole purpose of some short historical flashbacks to connect some aspects of this new version with its predecessors.

Some Basic Terms of Stereoscopic Vision and Film

Aware of the physiological foundations, Münsterberg proposed enhancing the perception of depth in motion pictures with the stereoscope, which was an influential medium in the nineteenth century. He even described the anaglyph technique of a red and green projection and the corresponding glasses in detail. "The effect is so striking that no one can overcome the feeling of depth under these conditions" (Münsterberg 1916: 49). Based on these thoughts, we might reflect on whether cinema requires "a perfect illusion of the outside world in sound, color, and relief," as Bazin ([1946] 1999: 201) noted in "The Myth of Total Cinema," or, in other words, whether the third dimension enhances the reality effect of cinematic representation.

When we consider the perceptual foundations of stereoscopy, it is intriguing to observe how closely the psycho-physical investigation of the senses and the emergence of a new technology were related, especially in the nineteenth century. The first construction of a working stereoscope is widely attributed to Charles Wheatstone, professor of experimental philosophy at King's College London, who described it in his *Contributions to the Physiology of Vision,* published in 1838. He built his stereoscope to explore the process of stereopsis, by which the eyes construct a depth impression by combining two different images projected onto the corresponding retinae.

A basic understanding of how stereoscopic representations make use of stereopsis in human vision is necessary to enable further discussion of aesthetics in stereoscopic films.[1] One fundamental of S3D is the *screen parallax*—the horizontal separation of the two images projected on the screen. It defines where an object is positioned in space. Negative parallax refers to a crossing of the corresponding lines of sight so that a point referring to the right eye is placed to the left of the corresponding point for the left eye. Negative or crossed parallax places objects in front of the screen (Figure 1) to produce the famous out-screen effect, which has often been regarded as one of the cheap gimmicks that make 3D cinema a fairground attraction. With positive screen parallax the corresponding points' position on the screen are placed with a difference such that the lines of sight meet behind the screen thus generating the in-screen effect in which objects appear to be placed behind the screen (Figure 2). With zero screen parallax objects appear to be positioned on the screen plane (Figure 3). For objects focused at infinity the two lines of vision have to be kept in parallel. Screen parallax is expressed in pixels, but it depends on the screen size. The bigger the screen, the smaller the parallax, because the effect of parallax is relative to the screen size in comparison to the distance between the eyes. Screen parallax in the positive range has a greater effect on depth perception because it occupies a larger angle of vision.

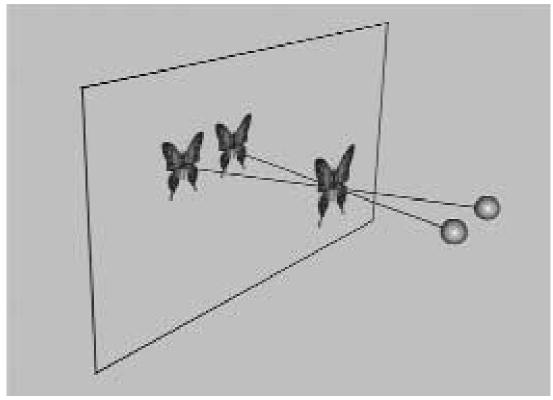

Figure 1. Negative screen parallax

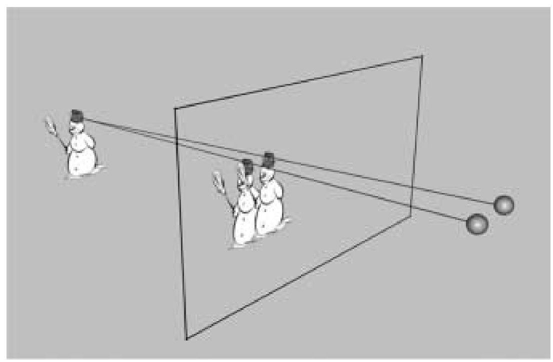

Figure 2. Positive screen parallax

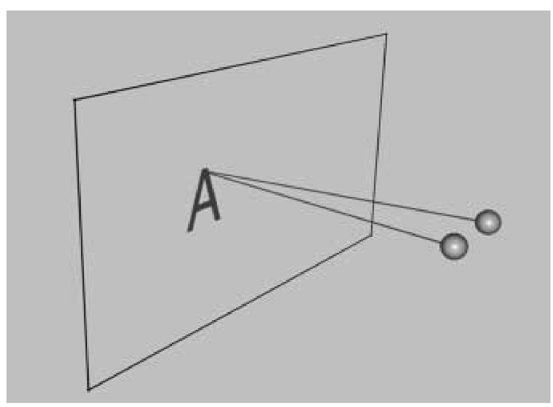

Figure 3. Zero screen parallax

Kluger, Josef (2009): *Visual Perception of 3D and Production Techniques.* Lecture at Hochschule für Fernsehen und Film HFF in Potsdam.

Another basic concept is horizontal *binocular disparity* (also called retinal or interocular disparity). It refers to the fixed distance between the two eyes. Multiple sources state that it is set at 63 to 65 mm on average as a standard for stereoscopic filmmaking. In real life this distance varies between viewers and is much smaller in children, a fact that has to be considered when stereoscopic movies are produced for a young audience (Luostarinen 2010).

When 3D movies are captured with two cameras this binocular disparity is mimicked by the so-called *interaxial disparity* between the two cameras. Interaxial distance (also called stereobase) has to be adjusted to the scene in front of the camera. The relationship between interaxial disparity and object distance transforms the depth and size of objects. If the ratio between interaxial distance and object distance is low, the objects look flattened; if it is high they are stretched. Equally important is the focal length of the camera lens. Lenses with long focal lengths compress the space in a scene. Based on these aspects—interaxial distance, object distance, and focal length in combination with parallax and screen size—a roundness factor can be calculated, which tells the stereographer whether the scene looks natural.

One of the foundations of a conflict between natural vision and stereoscopic film perception stems from the fundamental difference between the Euclidian space as projected onto the image plane by technical optics and the perception space as processed by the visual system. Or as Mendiburu (2009: 25) puts it more simply: Stereoscopic images do not exist as three-dimensional stimuli, but as a pair of flat images. This arrangement differs considerably from natural perception where the corresponding retinal points with zero disparity theoretically lie on what is called the *horopter*. In theory the horopter should form a circle called the Vieth-Mueller circle (Sedgwick 2005: 148). The empirical horopter, however, is flatter than the theoretical horopter, but in contrast to the projected images on the screen this area is not completely flat. There is a relatively large tolerance for points in space that can be fused to one three-dimensional perception by combining the different retinal images perceived by both eyes. This area is called *Panum's fusion area.*

Stereoscopic images do not exist as three-dimensional stimuli, but as a pair of flat images.

The final concept is the difference between accommodation—the contraction of the lens that defines where you focus on, and vergence—the synchronous movements of both eyes. In natural vision accommodation and vergence are always coupled. In other words we automatically focus on the objects we look at. This connection is known as the accommodation–convergence reflex or coupling. When perceiving a stereoptic projection we have to separate the two, and an accommodation-vergence conflict arises (Hoffman et al. 2008). Although we always focus on the screen we direct our glances to objects that are probably positioned in front of or behind the screen. According to most practitioners the disconnection between the two visual functions severely limits many creative choices established in two-dimensional motion pictures. Eye strain can be caused if these limitations are not taken into consideration. It should be noted that "the ability to control convergence and focus separately can be learned" (Mendiburu 2009: 21).

To some degree this practical insight contradicts a theoretical position presented by Cutting and Vishton (1995: 92ff.) who claim that vergence–accommodation coupling is limited to a rather small area in close vicinity to the observer. According to this view this conflict should not arise because the screen is positioned remote of this area.

Fundamental Aesthetic Problems of Stereoscopic Cinema

There are many well-known aesthetic problems associated with stereoscopic representations all of which call for specific arrangements of the profilmic scenery. I present some of these based on German film director Wim Wenders's documentary *Pina* (2010) on the world-famous Tanztheater (dance theatre) of Pina Bausch in Wuppertal, Germany.

In several interviews, Wenders claimed that this documentary would not have been possible without the new technology of digital stereoscopic cinema. He expressed his belief that only this technique can convey the concept of space and the corporeality of the dancers, both of which are central to this art form: "Die Körperlichkeit von Pinas Tänzern gibt es nur im Raum, die gibt es nicht als Abbild, nicht als Foto, nicht als Film" (The corporeality of Pina's dancers exists only in space. It does not exist as a representation, nor as a photo nor as a film. [*Translation by the author.*]) (Wenders 2011a: 25).

Though his arguments may seem plausible in theory, his film illustrates some of the fundamental problems of stereoscopic cinema in a striking manner. Many of the images would be very beautiful in 2D, but surprisingly show some of those shortcomings as textbook errors in S3D. These problems not only illustrate some fundamental aesthetic differences between representations in 2D and in S3D, but they also illustrate how artistic intentions by a director can clash with technical limitations. This clash is even more surprising

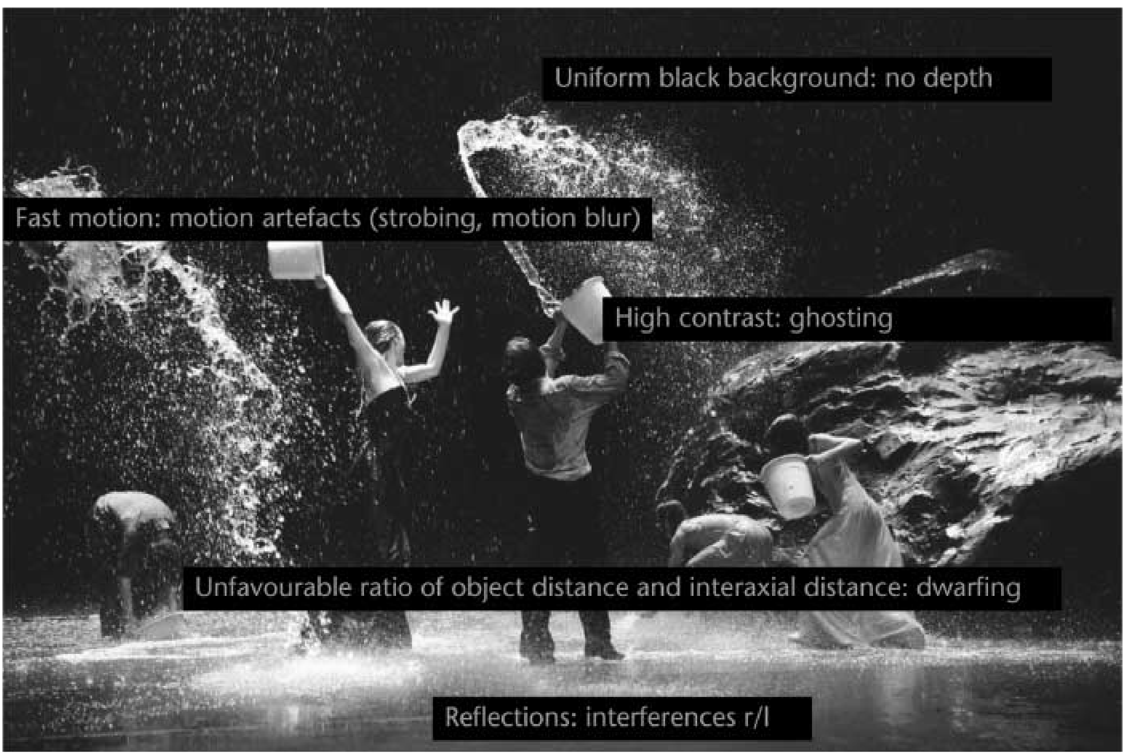

Figure 4. Still from Wim Wenders's Pina. PINA, Tänzer des Ensembles von "Vollmond" © NEUE ROAD MOVIES GmbH, photograph by Donata Wenders

in the case of a director like Wim Wenders who is widely regarded to be very technique-savvy.

First, most of the dance scenes in *Pina* are set on a theater stage in front of a uniform black background. When there is no structure presented in stereoscopic displays the eyes perceive no parallax and cannot calculate depth cues out of the stimulus. Similar problems based on a lack of structured detail occur when image parts are blurred, as in motion blur during fast motion. Additional motion artifacts stem from strobing caused by the shutter at a frame rate of 24 frames per second (fps). In contrast to 2D projections, in which every frame is projected twice to suppress the flicker, in stereoscopic projections the frames are projected alternately, albeit with a frequency of up to 144 fps, which is a triple projection of each frame, left and right in alternation. If the basic recording frequency is restricted to 24 fps, the missing parts between the images are too apparent and generate strobing. In their new productions both Peter Jackson with *The Hobbit: An Unexpected Journey* (2012), and James Cameron with *Avatar 2* (2014, projected) thus shoot at a frame rate of 48 fps. In the case of *Pina* these artifacts seriously undermine the corporeal impression by seeming to dissolve the extremities. It looks as though the bodies were losing their solidity or rigidity and becoming rubbery or even semi-fluid. Wenders (2011b) was aware of this problem: "Jede schnelle Armbewegung eines Tänzers produzierte den Eindruck, als sähe man für einen Bruchteil einer Sekunde zwei, drei oder vier Arme" (Every fast arm movement of a dancer produced the impression that he had two, three or four arms [*Translation by the author.*]). But there was no solution at hand. Oddly the fair skin tones in front of the black background intensify this effect and create ghosting due to the high contrast.

Another specific aesthetic problem of stereoscopic films is the rendition of reflections. Because they are based on the two cameras' positions in space and on their angle, the corresponding images of reflections differ significantly. Thus they are creating deviations that cannot be fused by the viewers or are perceived as strangely vibrating, semi-transparent surfaces that lie on top of the objects depicted. In addition to the conflict between natural and mediated stereoscopic perception often there is a technical reason that intensifies the problem. When so-called mirror rigs are applied in which the two cameras are positioned at an angle of 90° to each other, the semi-transparent mirror cuts out reflections by polarization and makes the rendition of reflections close to impossible thus severely limiting the possible array of profilmic arrangements since all reflecting and shiny materials are to be omitted. In *Pina* the resulting chatoyant effect was quite disturbing because it transformed the materiality of the depicted objects and environment. The floor of the stage is shiny, water is an important and recurring element, and large glass panels are visible in the background. I perceived the occurring interferences on these materials as affecting the very substance of the film's subject, the haptic quality in the interaction of bodies and matter as a central topic of Pina Bausch's dance theater.

In a similar fashion we are confronted with a perception problem specific to S3D, namely size perception. Dwarfism and gigantism are well-known problems of stereoscopic cinematography and they are owed to disproportionate interaxial disparity in relation to the distance of the objects depicted. In contrast to the real world where we experience size perception in relation to the position of our bodies, in film this relation obviously changes. In traditional montage with two-dimensional images we lack an important part of depth perception. Therefore we are much less sensitive to changes in size and accept more flexibility in the depiction of space. All the live action movies from *Avatar* (2009) to *Thor* (2011) show the problem of size rendition to varying degrees. In *Avatar* it is merely very difficult to develop a sense for the size of the Na'vi and the Avatars. In *Pina* the dancers often look like tiny figures on the stage, a problem Wenders was obviously aware of, because he makes deliberate self-reflexive use of it by combining a scene on the theater stage with a miniature of the set through an invisible cut.

Most of the issues described here are strictly tied to live action capture in S3D. In computer animation most factors can be controlled independently: movement can be planned in every aspect, reflections can be calculated on top of the renderings from the different angles (Gateau 2011), and size is completely independent from real-world constraints and can be adjusted to the specific needs for the scene.

Depth of Field in S3D

Depth of field is one specific area where we can observe conflicts between natural perception and stylistic conventions. Similar problems occur with editing pace, movement, image composition, lighting, motion blur, and so on. Many of them stem from a conflict between object and image perception, or between "fact" and "symbol," to draw again on Münsterberg's distinction. Depth of field is especially striking because the technological foundation is identical for both 2D and S3D. Thus the differences are perceptible only in the aesthetic domain where they evoke very different impressions in the viewers.

When I first met the new wave of stereoscopic cinema at the 2006 Association for Computing Machinery's Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH) conference in Los Angeles, the dominant approach was to capture everything with a large depth of field. In accordance with André Bazin's famous notion of deep focus, expressed in his essay on cinematic language, this paradigm presupposes that viewers should be able to explore the stereoscopic images at their own will. In his enthusiastic notes on the stereoscopic cinema, Sergei Eisenstein ([1948] 1999: 200ff.)

stresses this technology's ability to connect the foreground with the background in a striking manner.

One of the first of the new wave of S3D films, Robert Zemeckis's *Beowulf* (2007), was shot almost entirely with deep focus. In conjunction with overly sharp computer-generated imagery the pronounced deep focus created the impression of a sterile and utterly artificial world.

Although deep focus seems to be the first choice to make use of spatiality in S3D there are in fact serious limitations, one of them being the available space bracket or depth budget, as it is often called. Another limitation is the flattened look due to the cardboard effect caused by an unfavorable ratio between object distance and interaxial disparity. The depth budget results from several limitations that arise due to the conflict between natural perception and image perception (see Figure 5, which depicts the "comfort zone" in white. The depth budget is connected to this zone [Mendiburu 2009: 82])). One severe limitation is the disconnect between vergence and accommodation; the other limitation results from the ratio between inter-ocular disparity and screen size. Overall, negative and positive parallax should not surpass a certain percentage of the screen width set at a standard of 9 meters (30 feet). For smaller screens and depending on the viewing distance this might deviate significantly (see depth chart in Mendiburu 2009: 85). The positive parallax has to be set in relation to the native pixel parallax, that is the parallax equaling the interocular disparity. The positive parallax exceeding the double native parallax will cause the eyes to diverge, or to move outside the vertical axis. Divergence causes heavy eye strain if required for more than short moments, because we never diverge in our real life. According to Lenny Lipton, author of the comprehensive *Foundations of the Stereoscopic Cinema* (1982), divergence is possible under the following circumstances: "If the composition requires the viewer to observe the background in preference to the foreground, then divergence ought to be avoided. On the other hand, there are cases in which total divergence

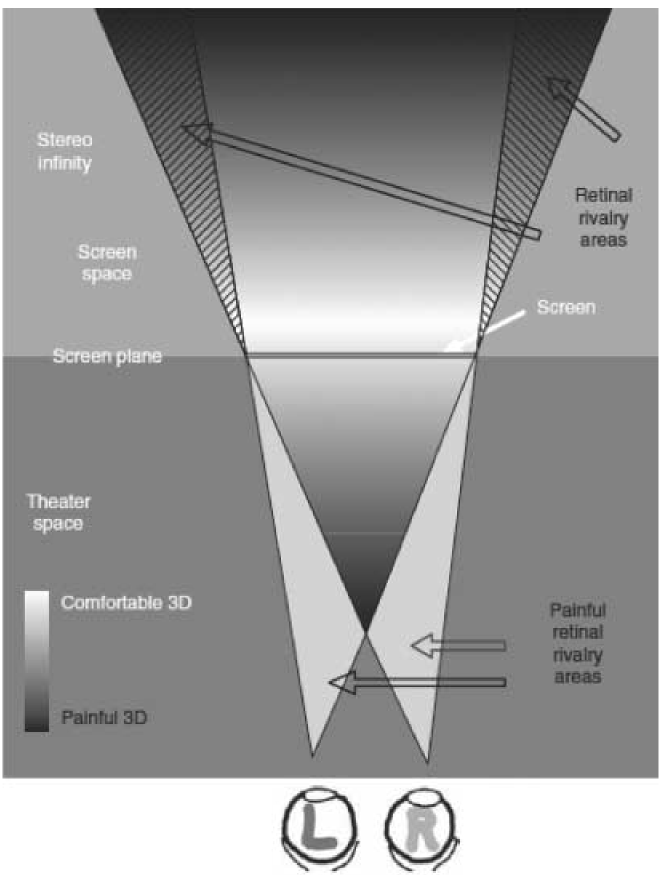

Figure 5. Comfort zone and corresponding depth budget Source: Mendiburu 2009: 82.

greater than 1° is permissible. For example, the background can be dark compared with the foreground" (1982: 192). Pete Kozachik (2009), the cinematographer of *Coraline* (2008), went further and proposed a maximum depth budget between -40 and +70 pixel separation, for short moments.

One possible reaction to the restricted depth budget is to place the scene in interiors with horizontally structured walls in the background as presented in the image from the German ballet film *Threesome 3D* (2010, see Figure 6). This film cleverly avoids the pitfalls discussed with the still from Wim Wenders's *Pina* and features no hard contrasts, no reflections, no fast movements. The horizontal structure on the wall supports the three-dimensional impression even in the background.

By looking at a stereoscopic version of Figure 7 you might have great difficulties in fusing the foreground and the background at the same time. Either you decide to converge—and it should be noted that you will have to converge off-axis, which can be painful on a big screen—and look at the foreground or you look at the background. Thus it is very difficult to build up the intended connection between Charlotte Rampling's character who looks out of the window and the group of street dancers waiting in front of the door. Furthermore the space depicted seems quite peculiar and distorted. Beyond the problems just noted, the still looks to have been put together out of different perspectives, which it most probably is. Some disorientation stems from the diagonal line of the window frame in the lower part of the image foreground. It might appear somehow counterintuitive that it can be much more difficult to perceive deep focus in S3D than in 2D because we would assume that depth in space should enhance the stereoscopic effect.

Figure 6. Still from Threesome 3D

Figure 7. Still from Streetdance 3D

The stop-motion animation in *Coraline* is still one of the most convincing examples of a staging that relies mostly on deep focus and works cleverly with the spatiality of the tiny world depicted. Beyond the observations on the narrative use of spatiality in this film as presented in *American Cinematographer* by Pete Kozachik (2009) and discussed by David Bordwell (2009) on his blog, *Coraline* is also a prime example of how the aesthetic features of stereoscopic images can be mastered in the best way possible. For me, it was no surprise that the author Lenny Lipton, quoted above, acted as a consultant for the stereoscopy in this film. In addition, as Kozachik (2009) points out, the very long shooting period of one and a half years provided a sound base of experience.

The camera team of *Coraline* used very small interaxials to depict the small universe, up to 3 mm for puppets in close-up and 3 to 10 mm for wide shots (Kozachik 2009). Therefore the stereoscopic effect is always gentle. Furthermore, Coraline's world is cluttered with many lovingly devised details. It offers a wealth of visual variety and the film displays the three-dimensionality of materials such as knitted clothes and the textiles the puppets are made of in a nicely fluffy and rounded manner that generates haptic impressions throughout the movie. Most of the images present a spatial continuum with floors and ceilings ranging from the foreground to the background. Exterior shots are often presented with a continuously ascending floor, often beautifully varied by many details such as flowers, plants, and Chinese lanterns. To enlarge the depth budget, Kozachik and his team changed the interaxial distance during camera movements continuously: "This allowed a deep 3-D effect at the wide end while making it easy for the audience to fuse left and right in the close-up" (Kozachik 2009).

One important factor for the convincing look of *Coraline* is the fact that the two corresponding images were shot with a single camera. In live action films many of the problems stem from the differences between the two physical cameras and lenses with regard to geometrical deviations, color rendition, and focal length. Even the finishing of the lenses is not standardized to the minute degree of detail that would be required for total congruence. These mismatches have to be corrected during the depth grading (Foundry Case Study 2010).

There is another difficulty in deep focus, namely the flattening of objects in the background. Because roundness is highly dependent on the ratio between interaxial distance and object distance, the farther objects are placed apart, the larger the stereobase should be. When this ratio is unfavorable for objects at a distance because it was calculated for objects at close proximity, the so-called cardboard effect can be observed. One approach to solving this problem is the layering of several shots with multi-rigs, each of which operates with an interaxial distance adapted to a certain range of the profilmic space (Kluger 2010). The resulting shots are then combined in compositing. This technique has also been applied to computer-animated movies such as *Tangled* (2010) (Neuman 2011).

A specific form of flattening occurs in wide shots of open landscapes when the background looks like a matte painting. The scenes do not open into infinity, but create a wall. However the stunning vistas of extended landscapes in the desert in *Inferno* (1953), directed by Roy Ward Baker and photographed by Lucien Ballard in Technicolor, are evidence that this problem can be overcome with sophisticated calculations of screen parallax. Also in the first full-length Russian stereoscopic feature film, *Robinzon Kruzo* (1946) directed by Aleksandr Andriyevsky and supervised by Semyon Ivanov (Drößler 2008: 13; Zone 2007: 168ff.) we can observe a very beautiful and convincing spatial arrangement in depth either by relying on continuous diagonal compositions in space or by layering several richly detailed planes. Most beautiful is a slow mystic cruise through lianas and between trees. Zone (2007: 169) quotes a contemporary review referring to this outstanding scene: "Out in the auditorium, about three rows in front of you, leaves and lianas materialize in the air, dangle and dance, and float away in Crusoe's face." This film was shot with a system similar to the Zeiss Ikon Raumfilm invented in the 1930s. It combined the two quadratic images on one strip by the implementation of a prism, in part on the then-new Agfacolor film stock (Drößler 2008: 13), in a way only made possible thanks to the auto-stereoscopic projection system described in Zone (2007: 167ff.). It is not surprising that Sergei Eisenstein (1948) wrote an enthusiastic article on "stereokino" after he had seen *Robinzon Kruzo,* which he regarded the best stereoscopic film he had ever seen.

In part to avoid the problem of extended depth, many wide shots in both *Avatar* and *Alice in Wonderland* (2010) are filled with haze, fog, or smoke in the background, thus limiting the depth range. In addition to serving this purpose, these background arrangements are also useful for computer-gener-

In part to avoid the problem of extended depth, many wide shots in both* Avatar *and* Alice in Wonderland *(2010) are filled with haze, fog or smoke in the background

ated imagery in general to help keep geometrical detail in very wide shots to a level still conducive to rendering.

In contrast to the wide acceptance of deep focus as an aesthetic choice, shallow depth of field in stereoscopic cinema has been the topic of a heated debate for years. In the interviews Ray Zone (2005) conducted with some twenty major proponents of S3D cinema from the 1950s on, a majority opted for deep focus. There are no hard and fast rules for the use of shallow focus in S3D, but there are certainly limits. Very shallow depth of field was applied in some shots of *Avatar* where only a part of a face is in focus. In these shots the blurred image parts dematerialize objects, causing the sharp portion of the face to look like a floating shell in S3D.

Shallow depth of field is nonetheless increasingly regarded as a viable choice in S3D. Some recent films, including *Thor* and *Harry Potter and the Deathly Hallows Part 2* (2011), have used it extensively. These films are conversions of films originally shot in 2D and as such the effect of all the parameters is very different—although in the case of *Thor,* the film was planned from the outset to be released with a stereoscopic version. While shallow depth of field is an important part of the cinematic language, it still needs to be adjusted to S3D, as *Alice in Wonderland* visual effects supervisor Ken Ralston (2010) noted. The film has some terrific shots, for instance one in which only a glass with the potion is in focus, beautifully lit, while Alice is a silhouette, barely perceptible when moving in the background.

There are several fundamental problems with restricted depth of field in stereoscopic images. First, it causes a conflict between object and image perception. When we look at objects in the real world, contour extraction is a necessary precondition to perceiving an object as an entity by providing the figure-ground separation (Peterson 2005: 171). When the depth of field is smaller than the depth of the visible part of the object, it dissolves the contours of the object by generating a visual continuum at the object's edge, creating an unusual impression oscillating somewhere between pictorial perception and object recognition. We perceive the stereoscopically projected elements as objects in space, but the dissolution of the contours renders them as semi-flat image parts. I want to add with regard to figure-ground separation that the common practice of accent contours in 2D films with back lighting in S3D stresses the cut-out effect by detaching objects from the background.

Second, every blurred element in a stereoscopic image poses a problem for the fusion of the corresponding image parts into a three-dimensional impression; this also applies to the blur in out-of-focus elements. As Robert Neuman (2011), stereoscopic supervisor for Walt Disney Animation Studios, has demonstrated in reference to *Tangled,* whether fusion is possible depends heavily on the structure of the blurred areas. He illustrated this fact with a variety of renderings of backgrounds, where we were able to see that contrast and a distinct pattern are a fundamental precondition for perceiving a spatial impression in these parts.

Third, a huge difference is made by whether the blurred area is in the foreground or in the background. Blurred backgrounds that are structured as proposed by Neuman are perceived as unobtrusive and natural. The situation differs considerably when objects in the foreground are out of focus, because then we tend to converge at these objects. In real life we do not converge without accommodating at the same time to see the objects we look at in focus. Therefore the glance at blurred objects gives rise to a feeling of frustration.

Fourth, in real life the blurred parts outside the foveal area of our visual field are only scanned to guide the gaze in case something is happening that requires further scrutiny. Therefore if large parts in a stereoscopic image are presented out of focus it results in a very unnatural perception as one can see in a displeasing shot from the famously failed conversion of *Clash of the Titans* (2010; see Figure 8). The effect is intensified when in addition to the large blurred image parts the object in focus is positioned at the side of the frame. In S3D this composition gives the impression of an unbalanced image.

Furthermore some images in *Avatar* were a special case of a blurred image due to two reasons. The image in Figure 9 was blurred in postproduction. It is the notoriously unpleasant Gaussian blur applied in compositing (Flueckiger 2008: 267). Gaussian blur is achieved by filtering sharp images through averaging by Gaussian distribution. In addition this image is the second part of a rack focus, where only a tiny computer-generated teardrop is in focus. Despite the very favorable reception of *Avatar*'s overall use of stereoscopy, James Cameron and his crew have been criticized for their use of shallow focus, especially when it is used in conjunction with convergence. If we analyze some of these images more closely, however, we may be able to see how elements

Figure 8. Still from Clash of the Titans

Figure 9. Blurred image in Avatar

of image composition and lighting contribute to the perception and provide a basis to differentiate the effects in more detail.

When we look back in history, we discover in earlier filmmaking an awareness about handling of restricted depth of field. In some of these early stereoscopic films from the 1950s as projected at the Filmmuseum in Munich during a presentation by Stefan Drößler, objects displayed in the foreground were often out of focus. However, they were placed mainly at the sides, thus providing a framing of the space depicted similar to curtains on either sides of a theater stage. Sometimes the objects were also very small or thin, such as branches or fine plants and ropes. In the German short film *6 Mädels rollen ins Wochenend (Six Girls Wheel into the Weekend)* (1939), a film that was, according to Drößler (2008: 11), never shown in public, we see one scene through foliage placed in the foreground. In a similar fashion there were icicles in a commercial for Volkswagen titled *Der weiße Traum (The White Dream)* (1950) and of course the notorious snowballs thrown toward the camera. To sum up these observations they clearly show how knowledgeable some of these early films were in applying shallow depth of field.

In *Avatar* there are blurred elements in the foreground that occupy extended areas in the frame. Especially in the scenes in the jungle they create a rather unfamiliar and quite disturbing perception of a gaze actively directed to an object out of focus. In Figure 10, almost everything in frame makes the audience converge on the lianas. Therefore they distract, not only by the placement of the object in the foreground but also by the glowing light—part, I must admit, of one of the film's most beautiful effects—created by the luminescent plants. In this shot these glowing lianas also move by way of a traveling shot around the embracing couple. Though one understands perfectly what is happening—thus the intrusion is not fundamental—we can still question this arrangement. Pete Kozachik (2009: 38) refers to such an

Figure 10. Glowing lianas in Avatar

arrangement as being tricky: "In 3-D, it's annoying to look through a tangle of soft-focus branches to see a sharp character, but if the branches are in another part of the frame, it seems to work." In his interview with Ray Zone (2005: 144), James Cameron talks about how he established a set of rules by analyzing stereoscopic movies, but then decided to break these rules to broaden his cinematic vocabulary. Even if, on inspection, certain shots in *Avatar* obviously violate some common rules, the unquestionably positive impact of this film's handling of stereoscopy by far exceeds the negatives mentioned here. Surely the most important asset of *Avatar* for its successful exploitation of stereoscopy lies in the dense and marvelous environment of Pandora whereby the stereoscopic rendition greatly enhanced the viewer's sense of presence in the fantastic world.

In a similar fashion but with completely different intentions, lighting and image composition can serve the purpose of suppressing legibility in order to deliberately disorient the spectator or hide information, in each case with the goal of heightening tension. In many shots from Josef von Sternberg's *Blonde Venus* (1932), for example, lighting and the arrangement of objects work against fast orientation and if we would track the eye movements on these images we most probably would get a vast distribution among different subjects. Many shots in the jungle in *Avatar* withdraw information as well by placing characters or animals behind plants where the blurred foreground objects are dark and thus do not generate conflicting targets of attention (Figure 11). These images thus function in much the same way as do correspon-

Figure 11. Supressed legibility in Avatar

ding shots in 2D cinema. From my reading of the manuals, however, I have the impression that many practitioners favor highly legible, evenly lit images in S3D, though my understanding is that this is not necessarily better suited.

Even more than shallow depth of field, rack focus has until recently been considered unsuitable in S3D. As a now conventionalized stylistic device to guide the viewer's attention it appears to have become more and more widespread. But it has a totally different effect in stereoscopic film than in two-dimensional movies. Due to the convergence-accommodation conflict the shifting focus is much more absorbing and can thus be more obtrusive.

In a rack focus from *Avatar* (Figures 12A and 12B) we see how it creates a conflict between image and object perception. Although on the image part we are drawn to the background, with the shifting focus our gaze tends to be directed to the foreground with the glowing animal and Neytiri's face. I would suggest that dialogue alone would have provided a similar shift of attention if both characters had been in focus during this shot.

Figure 12A and 12B. Rack-focus in Avatar

The situation is very different in a shot from *Alice in Wonderland* where the rack focus serves to tell an entire story, namely that we understand—a little bit ahead of Alice—that she has shrunk while the key lies on the table outside of her range of action. When the focus shifts to her face we know that she understands her situation and an alignment with her occurs creating a feeling of empathy. This rack focus is implemented in the best possible way because the two points of interest—the key and Alice's face—are located very close to each other. The key is a very small object and thus occupies a very small portion of the image in contrast to Neytiri in the example from *Avatar.*

Window Violation, Movement, and Visual Momentum

To close my overview on some aesthetic aspects of stereoscopic films I want to discuss some specific effects that arise from the spatial distribution of events and objects into the cinematic space—window violation, movement, and the resulting change in visual momentum.

Window violation occurs when an object overlaps the frame's border. Window violation is one of the most specific problems to occur in S3D. It leads to a breakdown of the stereoscopic effect due to retinal rivalry, that is conflicting monocular and binocular cues, because one eye lacks the corresponding information when depicted objects go beyond the frame's edges. In addition to the breakdown of the stereoscopic effect, the crossing of the frame's edges also draws attention to one of the notoriously critical zones in the cinema, because the frame's border also divides the world depicted in the film from our real world in the theater. It is thus a constant reminder that the space of illusion is only an illusion, and it threatens the audience's involvement in the film. In my experience as a S3D spectator window violations are highly disturbing when an object ranging from the center of the image to the foreground crosses the screen's edges as in a shot from *Alice in Wonderland* (Figure 13).

This effect also results from a conflict in image/object perception: images are always fragmentary, but objects only appear fragmented when occluded by other objects. Thus the eye perceives the arrangement in space as contradictory. While some objects seem to be positioned in front by negative parallax, they seem to be positioned as being behind by occlusion on the frame's border. As Bernard Mendiburu (2009: 182) points out there is the method of floating the window by masking a part on the side, thus eliminating the 2D/3D conflict. By floating the window it seems to be positioned in front of the screen, either parallel to it or also inclined to the actual screen (Neuman 2011). A problem remains, however, when objects in the foreground are fragmented by the edge and appear as though cut into pieces because the contour has been severely degraded. The effect and its questionable influence on

Figure 13. Window violation in Alice in Wonderland

the viewer's attention are even more pronounced when in pull backs objects suddenly pop into the frame.

Because the decision to render an S3D version of *The Polar Express* (2004) was made quite late in production the case is similar to the so-called death bed conversions (Coldewey and Wieland 2010), that is a failed production that was to be saved by converting it to S3D. Though for the computer-animated *Polar Express* it would have been possible to devise a better adapted framing, it seems that there was only a limited awareness of the problem at the time.

In S3D in general, the viewers' perception of the cinematic space is highly changed by movements, either movements in the diegetic space or camera movements and also movements from shot to shot through montage. In addition to stereopsis, motion parallax is the most important depth cue. If both are present as in stereoscopic cinema, they enhance each other, sometimes even to the point where they exceed the viewers' capability to process the overwhelming wealth of stimuli.

Slow and slightly curved lateral movements gently support the impression of depth by adding motion parallax. It seems that Alfred Hitchcock was well aware of this effect by implementing many traveling shots in his first and only stereoscopic movie, *Dial M for Murder* (1954). Fast lateral movements create the strobing motion artifacts mentioned in my discussion of *Pina*. For instance in *Harry Potter and the Deathly Hallows Part 2*, there is a fast circular movement, which I experienced as close to impossible to perceive. Pull backs often generate disorienting pop-up of lateral objects and thus strong window violations and distractors.

In contrast to lateral movements, forward movements seem best suited to creating a strong kinesthetic effect by taking full advantage of heightened depth cues in S3D. They are a key factor in delivering a pleasing and gripping cinema experience with stereoscopic formats. Thus those many ride scenes on the mountain banshees in *Avatar* and also the skydives of the Na'vi evoke an exciting feeling of joy; likewise, in a screening of *Rio* (2011), I heard children laughing with excitement as the birds flew over Rio de Janeiro. One of the strongest scenes in this regard in *Alice in Wonderland* is the flight over the treetops. In opposition to the staggered rendition of fast lateral movements, these movements in depth create the same perception as is present in our real-life experience of fast forward movements by the "transformation of the optic array . . . called optical flow field" (Sedgwick 2005: 142), namely the pattern where the optic array moves from a center outward. Therefore the situation supports the remark made by Anderson and Fisher Anderson (1980: 87) that we perceive real motion in the cinema: "It provides a perceptual basis, for instance, for Christian Metz's assertion that motion in the cinema is not a re-presentation, but a

In contrast to lateral movements, forward movements seem best suited to creating a strong kinesthetic effect by taking full advantage of heightened depth cues in S3D.

presentation, not the re-experience but the experience of motion." The more convincing experience of motion in forward movements than in lateral movements can be attributed to the fact that in this case the shift of the corresponding retinal points is relatively small from frame to frame. What is more, these images most often depict a figure in focus, moving in parallel to the (virtual) camera. Thus it is only the background that provides the optical flow field and this pattern is usually marked by heavy motion blur when the corresponding retinal points are distributed over a larger area.

To conclude these observations I draw on the concept of visual momentum proposed by Hochberg and Brooks (1978: 294). By visual momentum they mean a function of complexity and cutting rate that relates to the viewers' raised perceptual inquiry in processing stimuli with a correspondingly increased number of targets of attention. We could sum up this insight by concluding that the immersive sensory appeal of a film is the product of visual density and the cutting rate expressed as the reciprocal value of the average shot length. For S3D this formula should be expanded by the factors of depth, proximity, and movement, because each of these aspects greatly enhances the experienced visual momentum of the film. One reason for this effect is the time it takes to converge or even diverge based on changing stereoscopic depth information. Usually, filmmakers take this aspect into account by adjusting their cutting rate. A further strategy is the depth score or depth script that establishes the development of depth for a film in preproduction to set up a depth continuity. Its goal is primarily to match the depth cues to the narrative development by saving the most intense moments for dramaturgic nodal points. It also serves to avoid too many fast changes. Reducing the screen parallax before and after the cuts often softens transitions between shots. In many of the trailers for stereoscopic movies the visual momentum vastly exceeds the viewers' capacity, such as in *Step Up 3D* (2010). The same holds true for fast-cut fight scenes, such as those in *Avatar*, *Thor,* and *Harry Potter and the Deathly Hallows Part 2*. As Stefan Drößler (2008) points out, S3D cinema has always been caught in a paradox. Either it has made heavy use of S3D-specific effects and was criticized for its superficial play with spectacle, or it tried to rely more on classical storytelling and was blamed for not applying S3D in an appropriate way.

Thus there are two opposing paradigms, one claiming that S3D should not apply gimmicks but should first and foremost serve the story, the other claiming that S3D should enhance the viewers' experience in the cinema, sometimes even to the point of becoming an end in itself. Concluding from the many manuals and interviews I have read and the many presentations I have attended it is safe to say that most practitioners support the first paradigm.

There is a third possibility, one put into practice, for instance, by Alfred Hitchcock's *Dial M for Murder* and by Kenneth Branagh in *Thor*, despite the bad conversion of the latter. Style in these films is often deliberately exposed to create bodily effects of fear or vertigo. At the same time, these films add ironic distance to this exposition of style, especially in the moments of greatest tension as in the attempted murder in *Dial M for Murder,* in which the scissors in Grace Kelly's hand extend far into the theater.

Does stereoscopic cinema have a future? In 1948 Sergei Eisenstein was convinced that "To doubt that the 'space film' will be the standard tomorrow is as naïve as to doubt that there is a tomorrow at all" (Eisenstein 1948: 196). More recently, however, Walter Murch stated on Roger Ebert's blog in 2011: "3D doesn't work and never will. Case closed." Between these two extremes there is a compromise. With regard to the future development of S3D we can assume that stereoscopic cinema will survive in some genres, especially computer-animation and action-adventure including immersive science fiction movies and other body genres such as porn films.

Acknowledgement

I want to thank Maja Sánchez Ruiz for sharing the valuable information she researched for her master's thesis.

Barbara Flueckiger is professor of film studies at the University of Zurich. Her research focuses on the interaction between technology and aesthetics, especially in the digital domain. She has published books on sound design and on computer-generated imagery. Her current research project investigates the digitization of archival film. In fall 2011 she was a research fellow at Harvard University where she explored material and aesthetic aspects of color in film.

Notes

[1] For more information on technical aspects of stereoscopic films see Lipton (1982) and Mendiburu (2009) and Sedgwick (2005) on space perception.

References

Anderson, Joseph D., and Barbara Fisher Anderson. 1980. "Motion Perception in Motion Pictures." Pp. 76–95 in *The Cinematic Apparatus,* ed. Teresa de Lauretis and Stephen Heath. London: Macmillan.

Bazin, André. [1946] 1999. "The Myth of Total Cinema." Pp. 199–203 in *Film Theory and Criticism: Introductory Readings,* ed. Leo Braudy and Marshall Cohen, 5th ed. New York: Oxford University Press, 1999.

Bordwell, David. 2009. "Coraline, Cornered." http://www.davidbordwell.net/blog/2009/02/23/coraline-cornered (accessed 10 July 2011).

Coldewey, Georg, and Michael Wieland. 2010. "Converting Film into the 3rd Dimension." Paper presented at fmx conference, Stuttgart, 4 May.

Cutting, James E., and P.M. Vishton.1995. "Perceiving Layout and Knowing Distances: The Interaction, Relative Potency, and Contextual Use of Different Information About Depth." Pp. 69–117 in *Perception of Space and Motion,* ed. W. Epstein and S. Rogers. San Diego, CA: Academic Press.

Drößler, Stefan. 2008. "Kino in der dritten Dimension. Streifzug durch die Geschichte des 3D-Films." *Film-dienst* 1: 5–10, 2 : 9–13, 4: 44–47, and 5: 40–44.

Eisenstein, Sergei. "Über den Raumfilm." In: *Das dynamische Quadrat. Schriften zum Film,* translated by Oksana Bulgakova and Dietmar Hochmuth. Cologne: Röderberg, 1988.

Flueckiger, Barbara. 2008. *Visual Effects. Filmbilder aus dem Computer.* Marburg: Schüren.

Foundry Case Study. 2010. "Nuke and Ocula Deliver Stellar Performances on *Avatar.*" http://www.i3dpost.eu/dyn/1268068255879/CaseStudy_Avatar.pdf. Accessed July 18, 2011.

Gateau, Samuel. 2011. "Implementing Stereoscopic 3D in Your Applications." Presentation at fmx conference, Stuttgart, 5 May.

Hochberg, Julian, and Virginia Brooks. 1978. "Film Cutting and Visual Momentum." In Mary A. Peterson, Barbara Gillam, and H.A. Sedgwick. 2007. *In the Mind's Eye: Julian Hochberg on the Perception of Pictures, Films, and the World*. New York: Oxford University Press.

Hoffman, David M., Ahna R. Girshick, Kurt Akeley, and Martin S. Banks. 2008. "Vergence—Accommodation Conflicts Hinder Visual Performance and Cause Visual Fatigue." *Journal of Vision* 8(3): 1–30.

Kluger, Josef. 2010. "Grammar of the 3D Film Language." Paper presented at fmx conference, Stuttgart, 4 May.

Kozachik, Steve. 2009. "2 Worlds in 3 Dimensions." *American Cinematographer* 90 (2), http://www.theasc.com/ac_magazine/February2009/Coraline/page1.php (accessed 10 July 2011).

Lipton, Lenny. 1982. *Foundations of the Stereoscopic Cinema: A Study in Depth*. New York: Van Nostrand Reinhold.

Luostarinen, Tomi. 2010. "Demystifying 2D to 3D Conversions." Paper presented at fmx conference, Stuttgart, 4 May.

Mendiburu, Bernard. 2009. *3D Movie Making: Stereoscopic Digital Cinema from Script to Screen*. Oxford: Focal Press.

Münsterberg, Hugo. 1916. *The Photoplay: A Psychological Study.* New York and London: Appleton.

Neuman, Robert. 2011. “Stereoscopy From XY to Z.” Paper presented at fmx conference, Stuttgart, 5 May.
Peterson, Mary A. 2005. “Object Perception.” Pp. 168–203 in *Blackwell Handbook of Sensation and Perception,* ed. E. Bruce Goldstein. Malden, MA: Blackwell.
Ralston, Ken. 2010. “Alice in Stereoland.” Paper presented at fmx conference, Stuttgart, 4 May.
Sedgwick, H.A. 2005. “Visual Space Perception.” Pp. 129–167 in *Blackwell Handbook of Sensation and Perception,* ed. E. Bruce Goldstein. Malden, MA: Blackwell.
Wenders, Wim. 2011a. “Die 3-D-Technik ist wie für den Tanz gemacht.” *Filmbulletin* 53 (314): 25–26.
———. 2011b. “Filmisches Neuland”. http://www.pina-film.de/de/ueber-3D.html (accessed 28 July 2011).
Zone, Ray. 2005. *3D-Filmmakers. Conversations with Creators of Stereoscopic Motion Pictures.* Lanham, MD: Scarecrow Press.
———. 2007. *Stereoscopic Cinema and the Origins of 3D Film, 1838–1952.* Lexington: University Press of Kentucky.

Filmography

6 Mädels rollen ins Wochenend. Test film, shot in 1939, shown in 1941 in a closed screening only. Germany.
Andriyevsky, Aleksandr. 1946. *Robinzon Kruzo.* USSR.
Baker, Roy Ward. 1953. *Inferno.* USA.
Böhner-Film. 1952. *Der Weisse Traum.* West Germany.
Branagh, Kenneth. 2011. *Thor.* USA.
Burton, Tim. 2010. *Alice in Wonderland.* USA.
Cameron, James. 2009. *Avatar.* USA.
———. 2014 (projected). *Avatar 2.* USA.
Chu, Jon M. 2010. *Step Up 3D.* USA.
Giwa, Max, and Pasquini, Dania. 2010. *StreetDance 3D.* UK.
Greno, Nathan, and Howard, Byron. 2010. *Tangled.* USA.
Hitchcock, Alfred. 1954. *Dial M for Murder.* USA.
Jackson, Peter. 2012. *The Hobbit: An Unexpected Journey.* New Zealand and USA.
Louis Leterrier. 2010. *Clash of the Titans.* UK and USA.
Saldanha, Carlos. 2011. *Rio.* USA.
Selick, Henry. 2008. *Coraline.* USA.
Sternberg, Josef von. 1932. *Blonde Venus.* USA.
Vialkowitsch, Nikolai. 2010. *Threesome 3D.* Germany.
Wenders, Wim. 2010. *Pina.* Germany.
Yates, David. 2011. *Harry Potter and the Deathly Hallows Part 2.* USA.
Zemeckis, Robert. 2004. *The Polar Express.* USA.
———. 2007. *Beowulf.* USA.