Impacts of stereoscopic vision:

Basic rules for good 3D and avoidance of visual discomfort

Conflicts of depth cues, binocular rivalry and how to avoid it or "fix it in post"

No matter if "native" 3D or conversion from 2D to stereoscopic 3D, good 3D must not only have a high technical quality but also reduce visual discomfort by taking all aspects of the binocular vision system into account. This is essential for the consumer's 3D experience and thus for the success of the hardware- and the entertainment industries with this new format.

This report will initially review the human vision system in depth in order to understand why stereoscopic imagery can cause eye fatigue, headache or sickness. Then, the main conflicts of depth cues and binocular rivalry are described in detail. Finally, the last section addresses how these stereoscopic artefacts can be avoided or possibly fixed in a 3D live production ("native 3D"), a high-quality 2D-to-3D conversion and in a real-time conversion process.

1 The human vision system and the perception of depth

1.1 The monocular and binocular field of view

When capturing a 3D environment with a standard 2D camera, the depth information is lost. Humans have two eyes that capture their environment from two slightly different perspectives. The normal human eye distance ranges between 63 and 65 mm. The human brain processes the visual information and generates a stereoscopic depth perception.

The following figure shows a cross sectional area of the head with the visual cortex, the visual nerve system and the monocular and binocular field of view. The left monocular field of view is mapped to the right part of the visual cortex and vice versa. On the other hand, the left and right binocular fields of view, which are responsible for the stereoscopic depth perception, are mapped to the same side of the visual cortex.

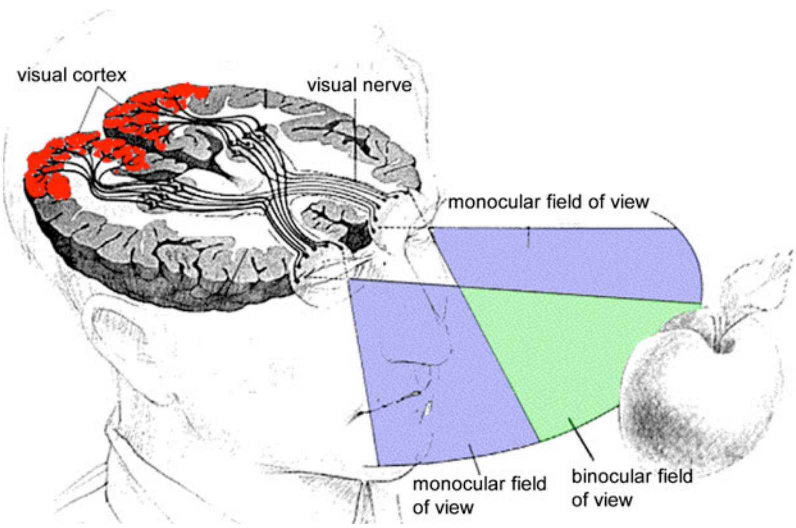


Figure 1: The human vision system

Author: Dr. Sebastian Knorr, imcube labs Berlin

1.2 Accommodation and convergence

Accommodation and convergence of the humans' eyes are needed to receive a clear image of the target object and to achieve single binocular vision in the visual cortex.

Accommodation is the process by which the vertebrate eye changes optical power to maintain a sharp image (focus) on an object as its distance changes (as depicted in Figure 2).

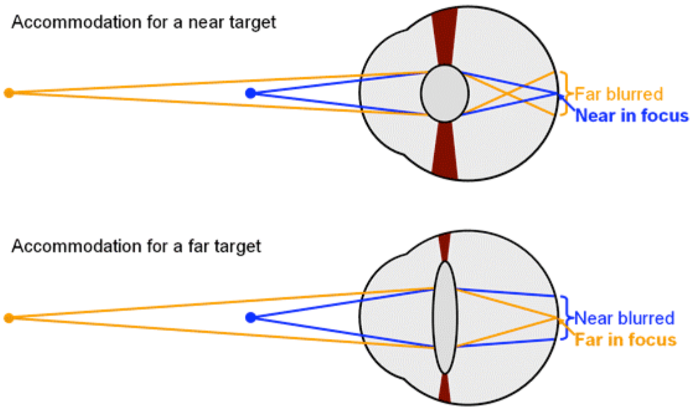


Figure 2: Accommodation

In ophthalmology, convergence is the simultaneous inward movement of both eyes towards each other, usually in an effort to maintain single binocular vision when viewing an object (as shown in Figure 3).

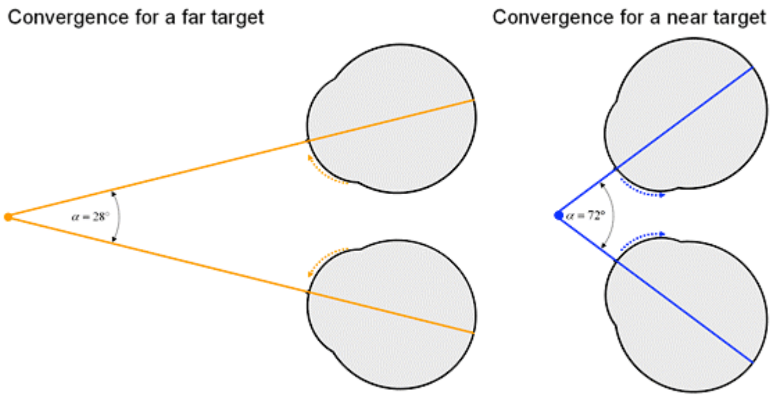


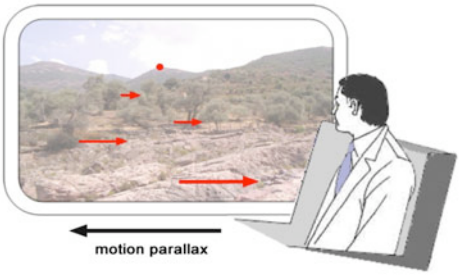
Figure 3: Convergence

1.3 Perception of depth

Depth perception arises from a variety of depth cues. These are typically classified into binocular cues that require input from both eyes and monocular cues that require the input from just one eye. A conflict of depth cues is the main cause for visual discomfort in stereoscopic vision. Thus, all depth cues are reviewed in detail.

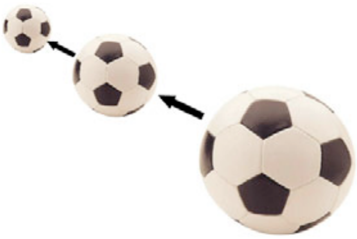
1.3.1 Monocular depth cues

Motion Parallax:



When an observer moves, the apparent relative motion of several stationary objects against a background gives hints about their relative distance. If information about the direction and velocity of movement is known, motion parallax can provide absolute depth information. This effect can be noticed when sitting in a train. Nearby things pass quickly, while far off objects appear stationary.

Depth from motion:



One form of depth from motion, kinetic depth perception, is determined by dynamically changing object size. As objects in motion become smaller, they appear to recede into the distance or move farther away; objects in motion that appear to be getting larger seem to be coming closer. Using kinetic depth perception enables the brain to calculate time to crash distance (aka time to collision or time to contact) at a particular velocity.

Light and shading:

The way that light falls on an object and reflects off its surfaces, and the shadows that are cast by objects provide an effective cue for the brain to determine the shape of objects and their position in space.

Relative size:

If two objects are known to be the same size (e.g., the soccer balls on the left) but their absolute size is unknown, relative size cues can provide information about the relative depth of the two objects. If one of the objects is larger, it seems to be closer than the others.

Familiar (absolute) size:

Since the visual angle of an object projected onto the retina decreases with distance, this information can be combined with previous knowledge of the object's size to determine the absolute depth of the object. For example, people are generally familiar with the size of their house. This prior knowledge can be combined with information about the angle it subtends on the retina to determine the absolute depth of the house in a scene.

Blur:

Selective image blurring is very commonly used in photographic and video for establishing the impression of depth. This can act as a monocular cue even when all other cues are removed. It may contribute to the depth perception in natural retinal images, because the depth of focus of the human eye is limited. In addition, there are several depth estimation algorithms based on focus and defocus (or blur).

Texture gradient:

Suppose you are standing in front of a field of sunflowers. The sunflowers which are closer can be clearly seen in terms of texture (shape, size and colour). As your vision shifts towards the distant field the texture cannot be clearly differentiated (the distant field looks like a homogeneous region).

Arial perspective:

Due to light scattering by the atmosphere, objects that are far away have lower luminance contrast and lower color saturation (e.g., distance mountains). In computer graphics, this is often called "distance fog" or "depth cuing". Objects differing only in their contrast with a background appear to be at different depths.

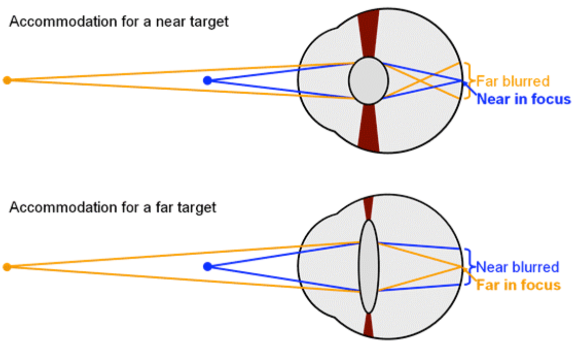
Perspective:

The property of parallel lines intersecting at infinity allows us to reconstruct the relative distance of two parts of an object, or of landscape features.

Interposition (occlusion):

Occlusion (blocking the sight) of objects by others is also a cue which provides information about relative distance. However, this information only allows the observer to create a "ranking" of relative nearness.

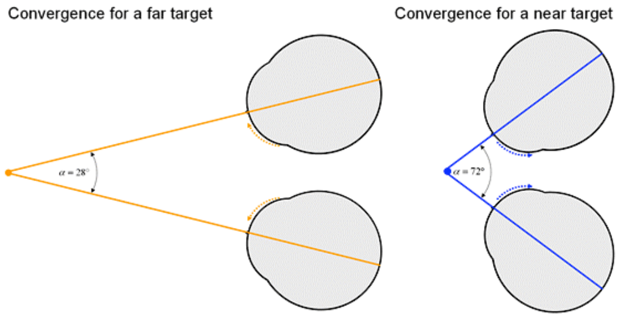
Accommodation:



When focusing on objects which have a large distance to the observer, the ciliary muscles stretch the eye lens, making it thinner, and hence changing the focal length. The kinesthetic sensation of the contracting and relaxing ciliary muscles (intraocular muscles) is sent to the visual cortex where it is used for interpreting distance (i.e., depth). Accommodation is only effective for distances less than 2 meters.

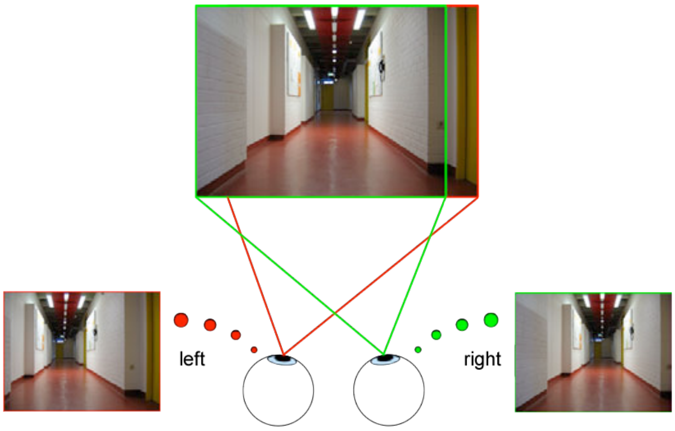
1.3.2 Binocular depth cues

Convergence:



This is a binocular oculomotor cue for depth perception. By virtue of stereopsis the two eye balls focus on the same object. In doing so they converge. The convergence will stretch the extraocular muscles. Kinesthetic sensations from these extraocular muscles also help in depth perception. The angle of convergence is smaller when the eye is fixating on far away objects. Convergence is effective for distances less than 10 meters.

Stereopsis or retinal disparity:



Animals or humans that have their eyes placed frontally can also use information derived from the different projection of objects onto each retina to judge depth. By using two images of the same scene obtained from slightly different angles, it is possible to triangulate the distance to an object with a high degree of accuracy. If an object is far away, the disparity of that image falling on both retinas will be small. If the object is close or near, the disparity will be large. It is stereopsis that tricks people into thinking they perceive depth when viewing Magic Eyes, Autostereograms, 3D movies and stereoscopic photos. For humans, stereopsis is effective for distances up to 30 or 40 meters.

2 Visual discomfort in stereoscopic vision

2.1 Conflict of depth cues

Conflicts of depth cues cause visual discomfort, which is the reason for headache and nausea. This is similar to the phenomenon "sea sickness". Often people are getting sick if the sea conditions are quite bad. The main reason is that the equilibrium organ detects the motion of the sea pretty well while standing on the deck of a ship. On the other hand, the visual system cannot detect this motion in the same way, i.e. there is a conflict of motion cues. The confusion between what the brain expects and what it sees can make people sick. Thomas P. Piantanida calls this phenomenon the *barfogenic zone*.

The same arises from a conflict of depth cues. However, not all conflicts of depth cues cause the same degree of visual discomfort. Depth cues have a different amount of intensity on the human visual system. Without going too much into detail, the most critical conflicts of depth cues are as follows:

1) Stereopsis vs. interposition:

   The most critical conflict of depth cues is the discrepancy between the binocular depth cue "stereopsis" and the monocular depth cue "interposition". If left and right views are displayed on a screen, the parallax of a pixel or object indicates whether it is perceived to be on the screen (no parallax), in front of the screen (negative parallax) or behind the screen (positive parallax) as depicted in Figure 4 (principles of stereoscopic depth perception). A conflict of depth cues occurs if for example object 1 is partly occluded by object 2 (i.e., object 1 must be behind object 2) but has more negative or less positive parallax than object 2 (i.e., object 1 is in front of object 2).

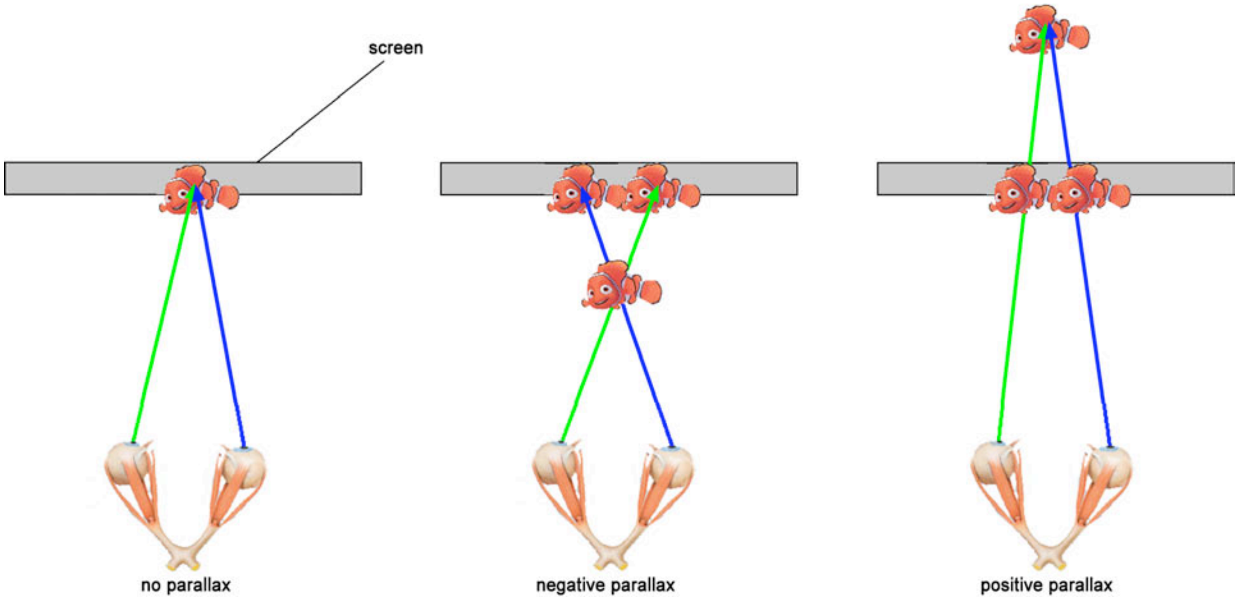


Figure 4: Stereopsis

A good example for this conflict of depth cues is *framing* or *frame violation*. This is a typical phenomenon when displaying S3D content on a screen with limited screen size. If an object has a negative parallax and crosses the image borders, it is truncated. The truncation indicates that the object must lie behind the screen (interposition). However, the negative parallax of the object indicates that the object is in front of the screen. Thus, a conflict of depth cues occurs as demonstrated in Figure 5.

Figure 5: Frame violation (left: negative parallax, right: interposition)

2) Stereopsis vs. perspective

Perspective is also a very strong depth cue. Rail tracks for example are parallel. In an image, which is a perspective mapping of the 3D environment onto a 2D plane, or on the retina the rail tracks seem to converge with increasing depth and intersect at the so called *point at infinity* as depicted in Figure 6.

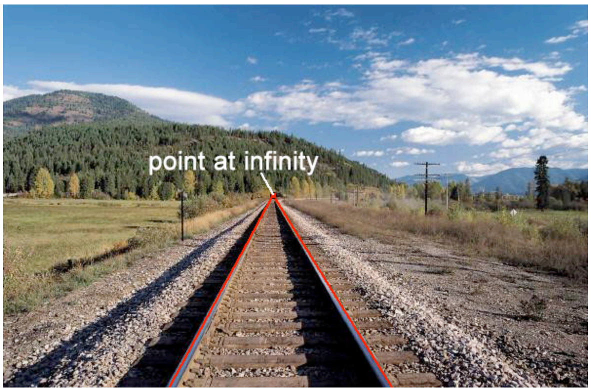


Figure 6: Intersection of parallel lines

If the point at infinity has the same parallax as the rail tracks on the bottom of the image, the human visual system would get confused heavily because of the conflict of depth cues.

3) Accommodation vs. convergence

   Besides *familiar object size*, accommodation and convergence provide absolute distance information. It is well known that, if a stereo image pair is displayed on a screen, the eyes solely focus on the screen. However, the stereopsis causes the eyes to converge to maintain single binocular vision when viewing an object. The distance information is dependent on the angle of convergence.

   Thus, there is a sensory conflict between the movement of the eyeballs (convergence) and the deformation of the lenses (accommodation).

4) Stereopsis vs. depth from motion

   As objects in motion become smaller, they appear to recede into the distance or move further away; objects in motion that appear to be getting larger seem to be coming closer. Thus, a conflict of depth cues occurs if an object becomes smaller but its negative parallax increases as depicted in Figure 7.

Figure 7: Example for stereopsis vs. depth from motion

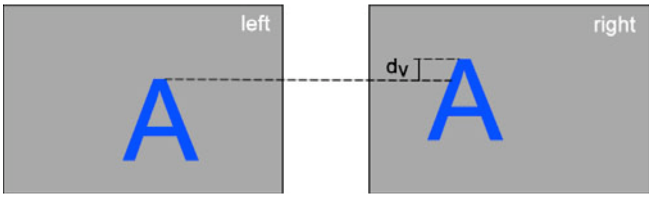
The ball, which is coming closer according to the binocular depth cue, gets smaller in size.

As already mentioned, there are several more combinations of depth cue conflicts. In general, these conflicts do not increase the visual discomfort to a great extent. Some conflicts of depth cues (e.g. relative size vs. familiar size) are even desired for creative reasons.

2.2 Binocular rivalry

Physical misalignments between the left and right stereo images lead to binocular rivalry, also called conflicting images. Binocular rivalry mostly occurs when shooting directly in 3D, i.e. with a stereo rig, or when displaying 3D with an imperfect projection system. All of the following misalignments may increase the visual discomfort dramatically and cause eyestrain, headache or nausea:

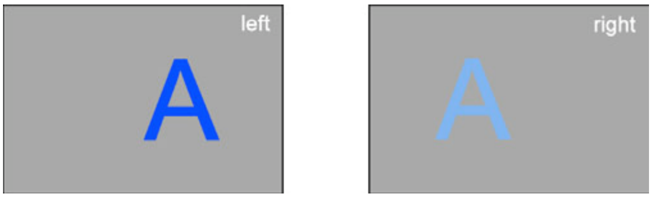
Vertical misalignment:



Characteristics:
Improper vertical alignment of left and right images

Reasons:
- Left and right camera or lens not properly matched
- Projection system not properly matched

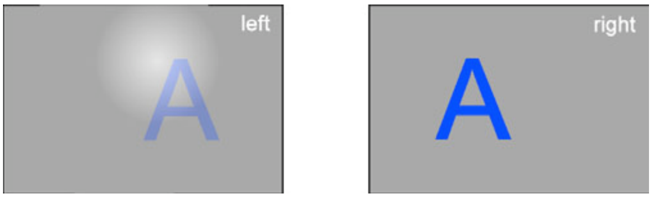
Luminance/ colorimetry:



Characteristics:
Left or right image is lighter or darker and/ or of different hue

Reasons:
- Left and right camera not properly matched
- Beam splitter diffraction

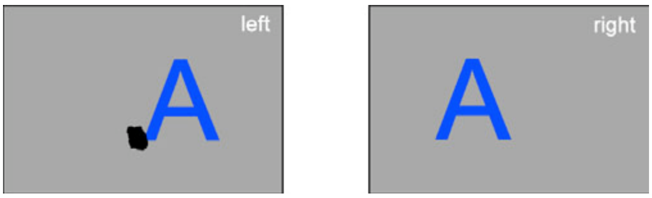
Reflections, Flares, Polarization:



Characteristics:
Reflections on objects not matching left and right view

Reasons:
- Camera positions/ angles
- Beam splitter polarization/ mirror rig

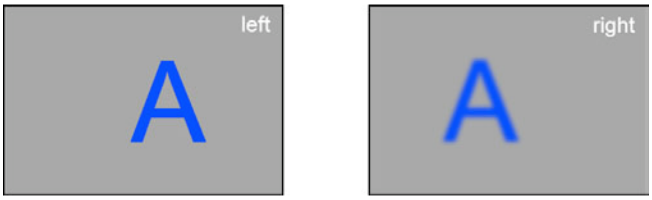
Contamination:



Characteristics:
Dust, water, dirt or other particles in one of the images

Reasons:
- Lenses or mirror not properly cleaned
- Bad environmental conditions

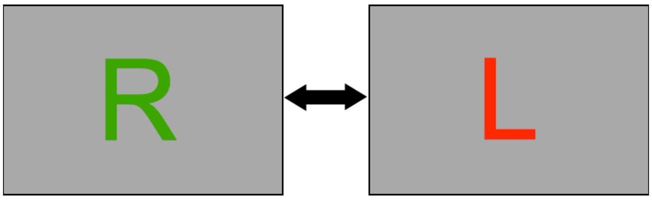
Depth of field:



Characteristics:
Focus of left and right camera not properly matched

Reasons:
- Different aperture settings
- Focal lenght of left and right camera not properly matched

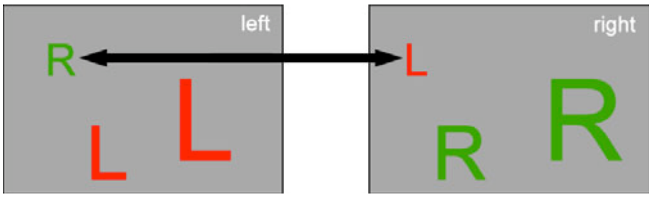
Pseudo 3D:



Characteristics:
Left and right image are swapped

Reasons:
- Camera cables are mixed up
- Left and right image are swapped when displaying on a 3D device
- Wrong naming of left and right view

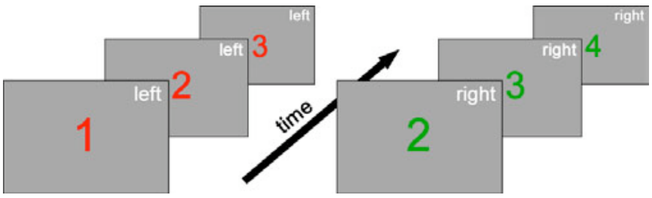
Partial pseudo 3D:



Characteristics:
Parts (layers) of left and right image are swapped

Reasons:
- Composition error in post

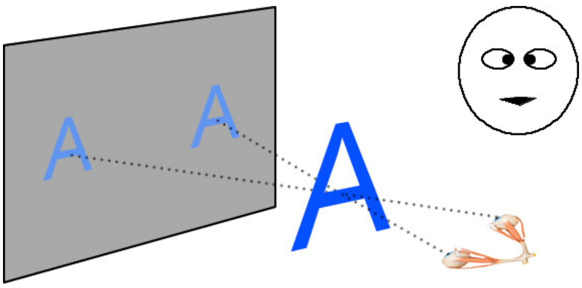
Synchronization:



Characteristics:
Left and right images are not properly synchronized

Reasons:
- Cameras are not synchronized
- Eddition errors in post

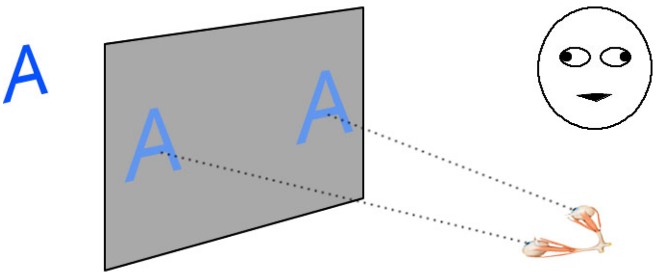
Hyperconvergence:



Characteristics:
Objects are too close to the viewer's eyes → single binocular vision not possible anymore

Reasons:
- Too much negative parallax
- Improper camera settings (e.g. too large baseline)

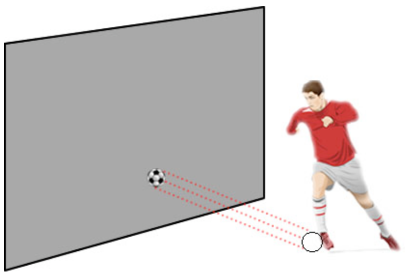
Hyperdivergence:



Characteristics:
Objects appear too far behind the screen → divergence of the viewer's eyes

Reasons:
- Two much positive parallax
- Improper camera settings

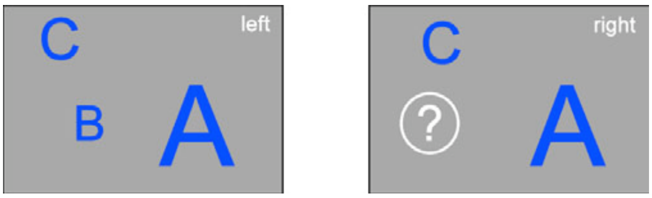
Depth mismatch:



Characteristics:
Improper depth composition (objects are in the wrong depth position)

Reasons:
- Composition error in post

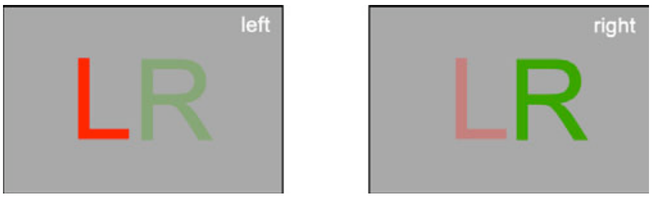
Visual mismatch:



Characteristics:
Objects that are just visible in one image

Reasons:
- Composition error in post

Ghosting:



Characteristics:
Double images (left image is slightly visible with the right eye and vice versa)

Reasons:
- Improper separation of left and right images by the 3D glasses
- Refresh rate of the 3D device

3 What makes good 3D

3.1 Basic rules for 3D productions

In any 3D project one has to take care of both the consistency of the depth cues and the avoidance of binocular rivalry. The following table summarizes how these conflicts can be avoided or maybe fixed.

Furthermore, it indicates the level of complexity if conflicts have to be fixed in post.

In Addition the table compares 3D live production with a stereo camera rig, high quality 2D-to-3D conversion and real-time conversion.

3D conflict	3D live production		High-quality 2D-to-3D conversion		Real-time conversion	
	How to fix	Complexity	How to fix	Complexity	How to fix	Complexity
Stereopsis vs. interposition	Only frame violation → floating window in post	low	- careful depth composition - floating window	middle (depending on # of layers)	- reduce depth budget - floating window - can only partly be fixed with complex image processing	high
Stereopsis vs. perspective	does not occur	-	moderate depth grading or 3D-modelling	middle	can only partly be fixed with complex image processing	high
Accommodation vs. convergence	reduction of depth budget (cannot be fixed entirely)	low	reduction of depth budget (cannot be fixed entirely)	low	reduction of depth budget (cannot be fixed entirely)	low
Stereopsis vs. depth from motion	does not occur	-	acurate depth composition/ modelling	low	can only partly be fixed with complex image processing	high
Vertical alignment	geometry realignment in post	middle	geometry realignment only for converging cameras	low	geometry realignment only for converging cameras	low
Luminance/ colorimetry	colour adjustment in post	middle	does not occur	-	does not occur	-
Reflections, Flares, Polarization	significant post-production work	high	does not occur	-	does not occur	-
Contamination	dust removal in post	high	does not occur	-	does not occur	-
Depth of field	significant post-production work	high	does not occur	-	does not occur	-
Pseudo 3D	swap left and right images	low	swap left and right images	low	swap left and right images	low
Partial pseudo 3D	swap incorrect layers in the composition	low	swap incorrect layers in the composition	low	cannot be fixed	cannot be fixed
Synchronization	significant post-production work	high	re-editing	low	does not occur	-
Hyperconvergence	- reduce negative parallax - reduce depth budget	middle	- reduce negative parallax - reduce depth budget	low	- reduce negative parallax - reduce depth budget	low
Hyperdivergence	- reduce positive parallax - reduce depth budget	middle	- reduce positive parallax - reduce depth budget	low	- reduce positive parallax - reduce depth budget	low
Depth mismatch	fix 3D composition	low	fix 3D composition	low	can only partly be fixed with complex image processing	high
Visual mismatch	fix 3D composition	low	fix 3D composition	low	does not occur	-
Ghosting	- reduce contrast - reduce depth budget - change convergence	middle	- reduce contrast - reduce depth budget - change convergence	middle	- reduce contrast - reduce depth budget - change convergence	middle

highest complexity (orange)
not completely fixable/ high visual discomfort (red)

3.2 Specific issues regarding stereoscopic 3D

The previous section described the basics regarding avoidance or fixing of conflicts of depth cues or binocular rivalry to reduce visual discomfort. However, there are additional issues that come along with the 3D content creation or the final displaying or projection of the 3D content. Those are:

- screen size and viewing distance
- shooting conditions on the set
- specific conversion problems

3.2.1 Screen size and viewing distance

It is well known that both the screen size and the distance of a viewer to the screen have an impact on the perceived depth. Moreover, the screen size may also have an influence on the occurrence of conflicts of depth cues and binocular rivalry, e.g. if stereoscopic content does not have any hyperdivergence when viewing on a 3D-TV display it may occur when viewing the same content on a large cinema screen.

Depth budget:

The depth budget is the distance between the nearest and furthest objects that the viewer perceives. The choice of the budget may be creative, but it has a significant impact on the visual comfort. In general, the depth budget relates to the maximum negative parallax and the maximum positive parallax which is measured in pixels. Many numbers about the maximum parallax in both directions are disseminated up to now. A critical limit is something around 3% of the horizontal resolution of an image (i.e. +/- 58 pixels for HD)[1]. But this approximate value is not always acceptable as indicated in the following example:

If the positive parallax of the furthest object is 58 pixels for HD resolution, its corresponding parallax in mm would be 302 mm on a 10 m cinema screen. Remember, the average human eye separation is between 63 mm and 65 mm. Thus, 3% positive parallax leads to hyperdivergence. On the other side, 58 pixels parallax for HD on a 50" display is equivalent to approximately 38 mm, which would not result in hyperdivergence.

The following table shows the critical limits assuming 3% parallax and additionally, taking the human interocular separation (approximately 64 mm) into account.

	2K resolution					
Screen size in m	max. neg. parallax			max. pos. parallax		
	in %	in pixel	in mm	in %	in pixel	in mm
1	-3,00%	-61,44	-30	3,00%	61,44	30
2			-60			60
4			-120	1,60%	32,77	64
6			-180	1,07%	21,85	
8			-240	0,80%	16,38	
10			-300	0,64%	13,11	
12			-360	0,53%	10,92	
14			-420	0,46%	9,36	
16			-480	0,40%	8,19	
18			-540	0,36%	7,28	
20			-600	0,32%	6,55	
22			-660	0,29%	5,96	
24			-720	0,27%	5,46	
26			-780	0,25%	5,04	

[1] This is often called the 1/30-rule of pixel parallax.

Screen size changes

The perceived depth of 3D content varies with the screen size. Basically, 3D content is produced for a fixed screen (usually a cinema screen). In general, shrinking the screen compresses the depth range whereas enlarging the screen expands the depth range, e.g. content that was originally created for a 50" 3D-display would result in an expansion of the depth range when displayed on a large cinema screen. Additionally, as indicated in the previous table, hyperdivergence could occur when using up to 3% positive parallax.

The basic formular for the perceived depth *d* of an object in relation to the object separation *s* on the screen (parallax), the distance to the screen *D* and the eye separation *e* is:

$$d = \frac{e * D}{e - s}$$

The following figures demonstrate the effects of display size changes using the same parallax settings in pixels:

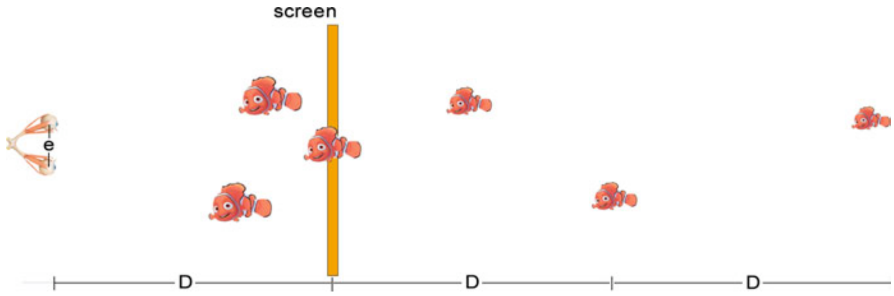


Figure 8: Perceived depth on the original screen

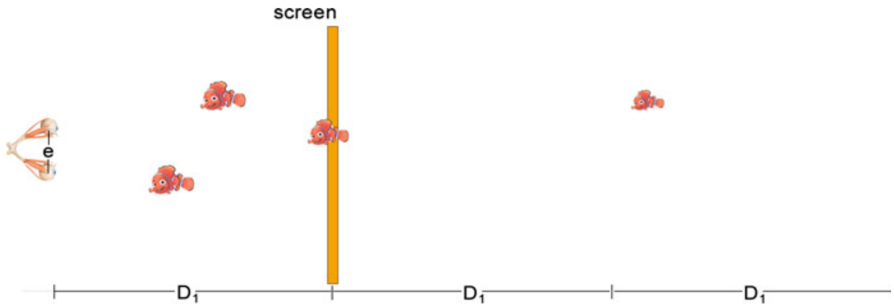


Figure 9: Perceived depth on a screen of double size

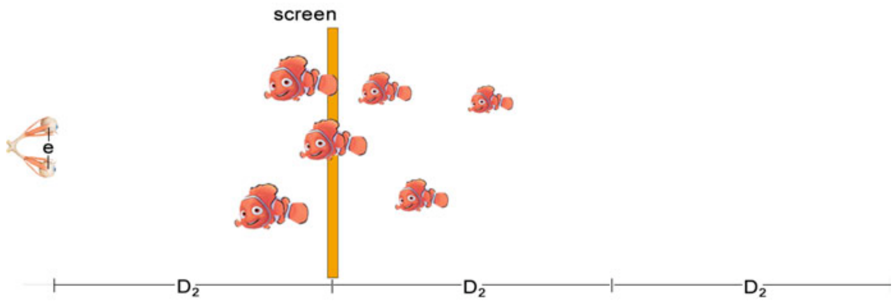


Figure 10: Perceived depth on a screen of half size

Actually, 3D content that was produced and optimized for a large screen can be displayed on a small screen without visual discomfort. However, the perceived depth is less intensive as depicted in Figure 10. On the other hand, 3D content that was produced and optimized for a small screen may cause visual discomfort because of the occurrence of hyperdivergence. The perceived depth is much more intensive on a large screen as demonstrated in Figure 9.

The latter is still an issue during 3D live production. Most of the 3D footage is produced for cinema screens. However, due to practical aspects the director and stereographer do just have small preview screens to monitor the captured 3D content during the shooting. Thus, it needs some experience to know how the 3D footage may look on a large screen.

Finally, framing effects the human visual system differently depending on the screen size. Since the borders of a small 3D-TV display are always within the viewing angle of the human's eyes, framing (interposition of objects and image borders) appears more drastically than on a large cinema screen were the borders might be outside the binocular field of view. Thus, the negative parallax should be reduced for smaller screen sizes.

Viewing distance

The previous formula also indicates that the perceived depth increases proportionally with the viewing distance to the screen as depicted in the following figure.

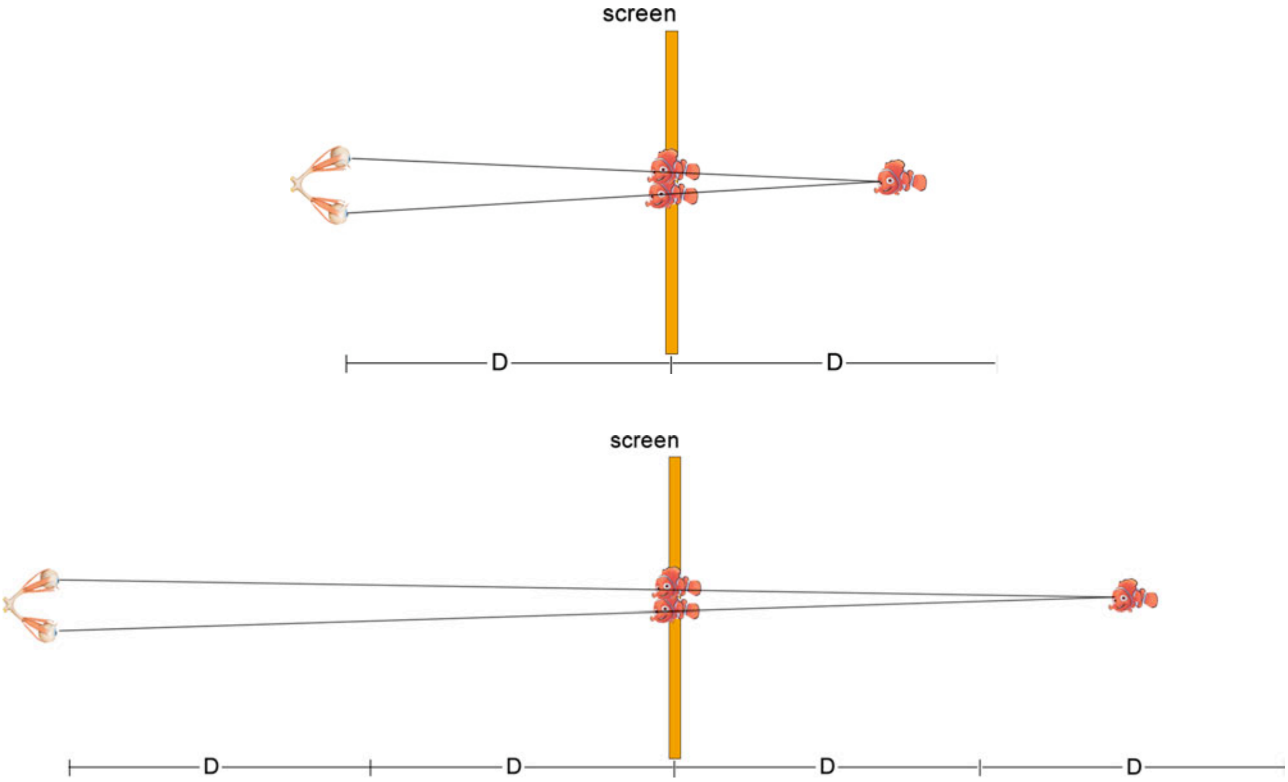


Figure 11: Perceived depth depending on the viewing distance to the screen

3.2.2 Shooting conditions on the set

As already indicated in section 2.2, binocular rivalry mostly occurs while shooting with a stereo camera rig. This section should give a brief overview of some practical issues regarding 3D live production.

First of all it is clear that the two stereo cameras are physically not identical due to construction tolerances. There are small variations e.g. of the two lenses, sensors, etc. Additionally, synchronization, alignment, zooming and converging of the two cameras are issues that are difficult to handle. There are some conditions on the set where the shooting with a stereo rig is much to complex or even impossible. When capturing e.g. a landscape from a helicopter's point of view, the baseline between the two cameras needs to be increased drastically to achieve a desired depth effect. It is either not possible to increase the baseline sufficiently or the previous mentioned issues like alignment or synchronization are even more difficult to handle.

Mirror-rigs that are widely used especially for close-ups, produce errors like different polarization, reflections, flares and contamination by nature. These errors have to be fixed in the post. The worse the shooting conditions the more likely is the occurrence of binocular rivalry. While shooting e.g. in a dessert, it is quite hard or almost impossible to keep the mirror clean since there are lots of sand particles in the air that even may scratch the mirror.

3.2.3 Specific 2D-to-3D conversion problems

Although binocular rivalry is not the main issue when converting footage from 2D to 3D, there are many other issues where conversion is quite complex and extremely difficult. These are shots that contain semi-transparencies, reflections, water, rain, fire and explosions.

The masking of raindrops during a rotoscoping or keying process for example is extremely labour-intensive if done moderately. Furthermore, accurate 3D modelling of the masked rain is almost impossible. Thus, it is more efficient to re-produce rain with vfx tools and add it to the converted stereo content.

3.3 Conclusion

3.3.1 3D live production

It is a fatal mistake to believe that shooting directly with a stereo camera rig is easier, less expensive and causes fewer errors. In fact, artefacts resulting from shooting with stereo camera rigs can be greater than those created throughout a 2D-to-3D conversion process.

In a stereo mirror rig for example, one camera has to focus through a beam splitter mirror, i.e. capturing polarized light, while the second camera is capturing the reflection of the mirror. This shooting results in binocular rivalry, namely: luminance differences, polarization, reflections, flares, contaminations, misalignments, etc.

Additionally, it is the nature of any 3D live shooting that different depth of fields, vertical misalignments or synchronization errors occur. In general, most of the errors have to be fixed in post. Due to the absence of compositing layers, post-production can be very complex and expensive.

3.3.2 Real-time 2D-to-3D conversion

Real-time conversion is currently a hot topic in the entertainment industry, especially in the 3D-display market. There is a great interest in selling the newest 3D display technology for the consumer market. Due to a lack of 3D content, real-time conversion seems to be the most cost effective and fastest way to enter this market.

Nevertheless, most experts believe that real-time conversion is just a gimmick due to the bad quality. It is impossible to watch an entire 90 minutes feature film without any such issues. Furthermore, real-time conversion does not have any creative potential which is essential for the movie industry.

As depicted in the previous table, not all depth cue conflicts and binocular rivalry issues can be solved properly. Thus, eye fatigue, headache and sickness is inevitable.

3.3.3 High-quality 2D-to-3D conversion

All issues regarding depth cues and binocular rivalry can be managed in a high-quality conversion process if done properly. However, high-quality conversion is labour intensive and much more expensive than real-time conversion. As a result visual discomfort, which is essential for good 3D, can be avoided.

Compared to 3D live production and post-production, 2D-to-3D conversion is mostly less expensive. Of course, native 3D can be produced more accurately with stereo camera rigs if done properly. However, to avoid or reduce artifacts significant efforts have to be made in post-production.

3.3.4 3D animation (CGI)

3D animation was not addressed in the previous table because native 3D with no artifacts can be achieved quite easily. In fact, production costs are more a matter of workflow optimization. However, in some CGI-productions 2D-to-2D conversion is used in the compositing process because of cost efficiency. If the render layers are available, the labor intensive processes of rotoscoping and clean-plate creation can be reduced substantially.

4 Displaying stereoscopic content

Besides the proper production of 3D content, the displaying conditions may cause visual discomfort as well. Since many techniques like anaglyph, polarization (e.g. RealD), shutter (e.g. Xpand), wavelength multiplexing (e.g. Dolby 3D), lenticular lenses, parallax brarrier, etc. exist on the market, the underlying technology needs to be reviewed to understand their impact on the human visual system. However, this is out of the scope of this paper. A detailed description will be part of an upcoming "Technical Information" about this special topic.

Dr. Sebastian Knorr is one of the founders, managing director and head of the R&D and Knowledge Management department of imcube labs GmbH, Berlin (www.imcube.com). Since 2002 he worked as researcher in the field of 3D image processing and computer vision in the Communication Systems Lab at Technische Universität Berlin, Germany.

In 2007 Dr. Knorr invented the patented process of automatic 2D to 3D image conversion with highly advanced computer vision technology. He received the German Multimedia Business Award of the Federal Ministry of Economics and Technology in 2008.

imcube labs is an R&D and Software Development company which is licensing its software to production companies worldwide. It has developed “imcube cinema”, a specialized 2D-to-3D conversion framework and compositing software that takes care of all the depth cues and binocular rivalry issues. It integrates many fundamental features that are needed for an optimal workflow with flexible user-interactivity and creativity: depth assignment and compositing, layer management, pre-visualization and screen-size adjustment. “imcube cinema” has been designed for the 2D-to-3D conversion of feature films and commercials for theatrical exhibition. “imcube home” for Blu-ray applications and “imcube tv” for the special requirements of 2D-to-3D conversion of broadcast content are under development.