



## Daniel Wagner

[+ Follow](#)

Heading Snap's Spectacles software development including OS, Computer Vision, Hand Tracking, Application and Mobile.

Published Aug 7, 2019

### [Daniel Wagner, Louahab Noui, Adrian Stannard](#)

Mobile Augmented Reality (AR) has become a hot topic. When reading popular media about recently released Head Mounted Displays (HMDs) such as Microsoft's "Hololens", Magic Leap's "Creator Edition" or DAQRI's "Smart Glasses" it seems AR is finally becoming ready for the masses. While these devices still have many limitations, the one component that is probably most limiting today is the display.

As a result lots of speculations and "feature requests" have recently been promoted by popular media. The common consensus seems to be that today's AR displays are great, if not for the field of view (FOV) that is still too small, which however will soon be great as well. Some display makers such as waveguide designer DigiLens also jumped on the FOV wagon and recently claimed that their AR displays will soon reach 150° FOV.

In this article we will try to explain why FOV is unlikely to expand drastically any time soon, but even more importantly we aim to point out that FOV is just one of about two dozen parameters that are important for AR displays. Neither are those other properties less relevant than FOV nor have they been solved already or can be resolved independently. In fact most AR display technologies today have been around for years in military or industrial scenarios. Recent progress was less in developing new break-through technology, but rather in bringing cost down.

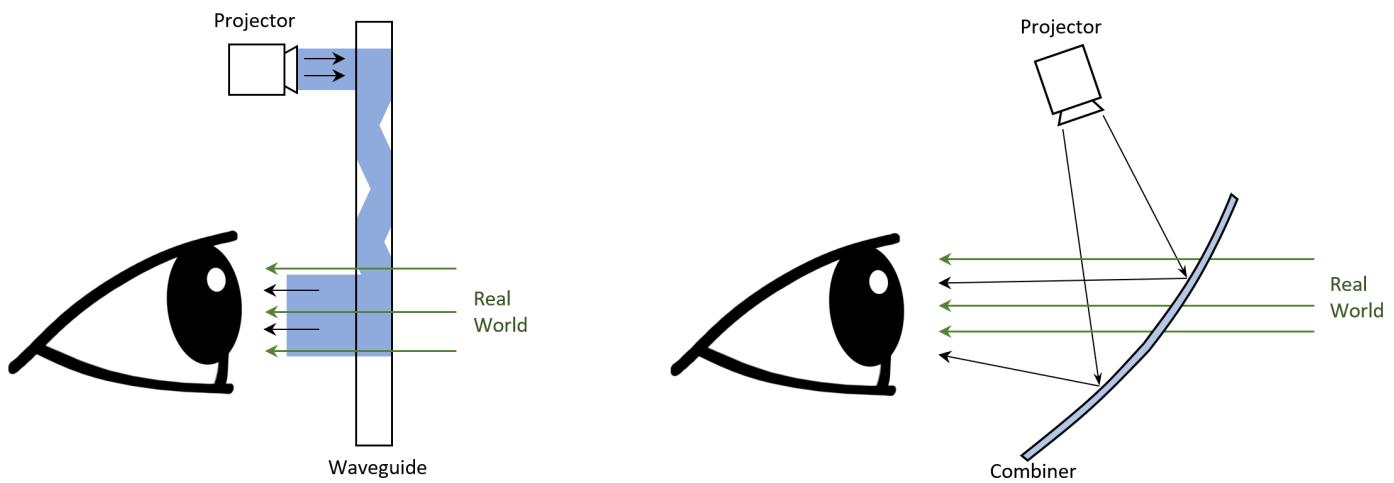
## Types of AR Displays

There are two main technologies for AR displays: optical see-through and video see-through. Both options have been explored in the past, but due to significant



Optical see-through displays allow the user to see the real world “directly” (through a set of optical elements though). These AR displays add virtual content by adding additional light on top of the light coming in from the real-world. Since this is an additive operation it is impossible to show black color or to darken the real world. Hence showing black virtual content or drawing shadows is not possible with today’s passive optics. There are experimental devices that can block incoming light on a per-pixel level, but these are far away from practical usage and are therefore not considered here.

Today there are two main types of optical see-through displays: waveguides and free-space systems (combiners). **Waveguides** (see left image in the figure above) are clearly dominant in the high-end space today and all the aforementioned devices are waveguide-based: A projector inserts an image at one location (in-coupling) of the waveguide. Due to total internal reflection (TIR) it bounces inside the waveguide (like in a fibre optic cable) and then exits at another location (out-coupling) towards the eyes. Waveguides are popular, because they allow for elegant flat designs, but they come with a lot of (image quality) problems.



**Figure:** Optical see-through displays using a waveguide (left) and a free-space system (right)

The most popular alternative to waveguides today are **free-space systems**, which - due to using free-form elements - are complex to design, but optically still much

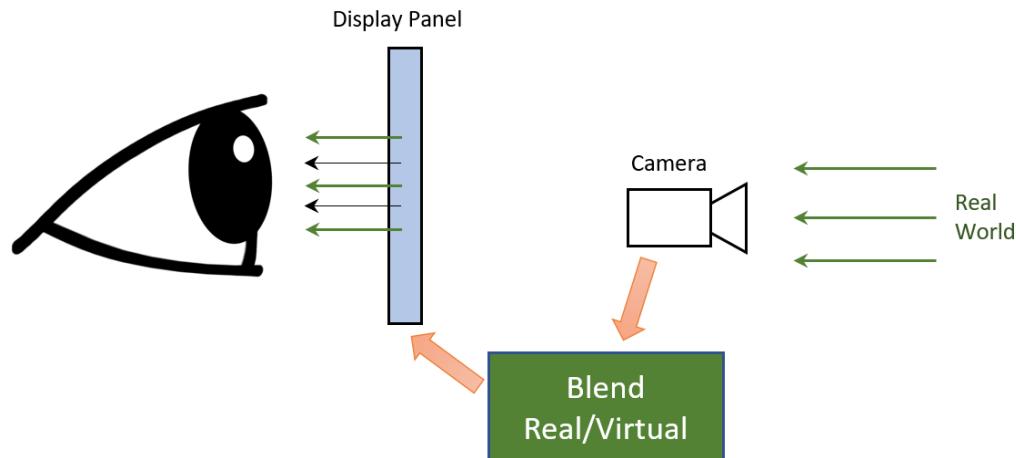


views.

Waveguides have a low optical efficiency and hence require powerful projectors. Today this is achieved by combining LEDs as light source with LCOS as image modulators. Free-space systems are usually efficient enough for using OLEDs, which are self-emissive and enable smaller projector designs.

### Video see-through displays

In the case of video see-through displays a pair of cameras record the real-world view, which is then shown on an opaque display, such as OLED or LCD. Virtual content is added using classic video mixing techniques, which means that any kind of operation including showing black virtual content and darkening the real world is possible.



**Figure:** Principle of a video see-through display

The ability of doing proper video mixing is definitely an advantage of video see-through over optical see-through. Still, basically all AR devices today use optical see-through and the reason is simple: In the case of video see-through all the challenges that we discuss in this article apply to the real world view and to virtual content. In contrast, in the case of optical see-through they only apply to virtual content, which can be much better controlled with clever UI design. Dynamic range is an obvious example here: While the human eye with its enormous dynamic range can clearly see a person in direct sunlight right next to a person standing in the shadows, today's



are also safety concerns as well as human factors which need to be addressed. Therefore, in the remainder of this article we'll only consider optical see-through displays.

## Design Parameters

In this article we postulate that field of view is just one of many essential design parameters of an AR display. Most of these design parameters are equally important as FOV. In the following we describe these most important properties, which are

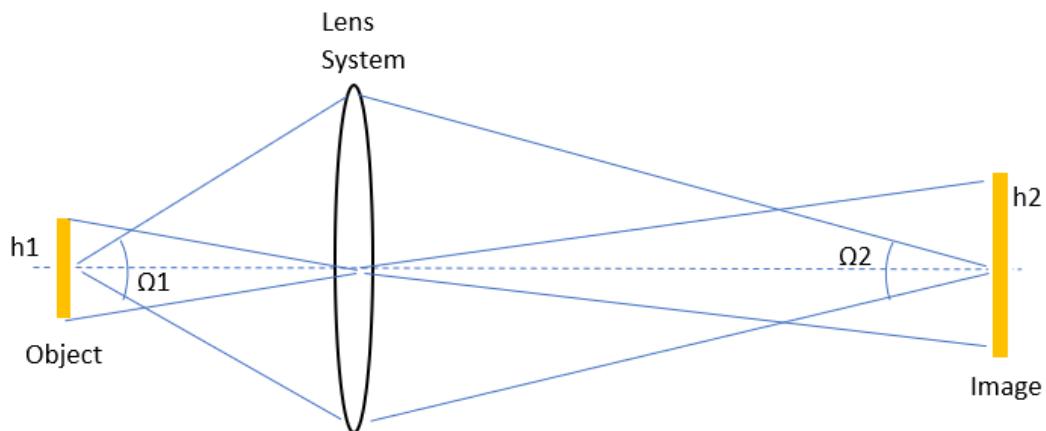
- Field of view
- Eye box size
- Brightness, transparency and duty time
- Contrast
- Uniformity & color quality
- Resolution
- Real-world distortions
- Virtual image distortions
- Eye safety
- Eye relief
- Peripheral vision
- Chromatic aberrations
- Depth perception



- Stray light

It is relatively easy to improve one parameter at the cost of others. For instance, increasing the field of view is not that hard if bulkiness and a small eye box are acceptable. However, most users would probably not want to use such a device. On the other hand, the combination of a large field of view and a large eye box in a small display is indeed highly challenging. Similarly, a larger eye box requires more light to achieve the same perceived brightness, so a more powerful light source is required.

Most of the parameters described in the following are not at their desired levels yet today. Hence, device makers aim at improving upon all of them. However, as just mentioned improving even just one parameter without sacrificing others is already difficult. The main reason for the trade-off is captured by the so-called **etendue** of the system, which is a geometric invariant defined as in the formula below. Etendue must be conserved, analogous to the conservation of energy. In its simplest form it states that the product of solid-angle of light with surface area for a given light source must be a constant (a very friendly introduction to etendue is given in [this illustrated piece by xkcd](#)).



**Figure:** Rules of etendue: As  $h_2$  increases  $\Omega_2$  must decrease.

In the image, the object of height  $h_1$  functions as light source. For the geometry of the lens system shown, the resulting image height  $h_2 > h_1$  is magnified, while the



$$G = n^2 \cos \theta dA d\Omega$$

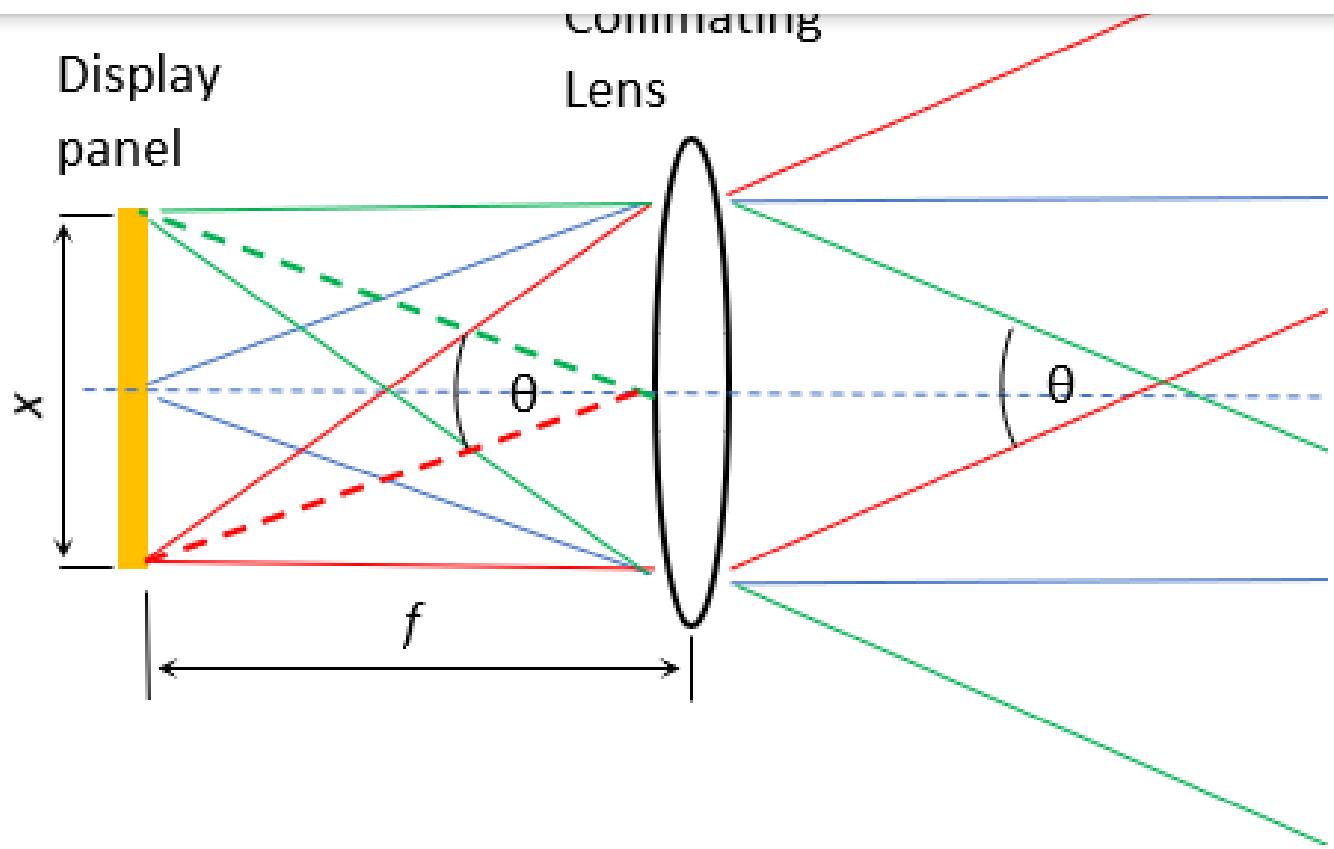
Where  $n$  is the refractive index of the medium,  $\Theta$  is the angle of the emitted (or received) cone of rays of area  $dA$ . We note there is no standard symbol in use for etendue, however “G” and “dG” is used frequently within the optics community. An analogous term appears in the paraxial limit known as the Lagrange invariant, which states:

$$h_1 u_1 = h_2 u_2$$

Where  $h_1$  and  $h_2$  are the object and image heights as before, and  $u_1$  and  $u_2$  are the object and image ray angles respectively. An alternative expression for G is sometimes used for working with microscope objectives with NA given by:

$$G = \pi A (NA)^2$$

Etendue has an impact on the efficiency of light engine design and projector size if trying to expand the exit pupil for a constant field-of-view. Consider for instance a simple projector design in the image above in which a collimating lens of focal length  $f$  collimates a micro-display panel of width  $x$ , (we shall limit ourselves to one



**Figure:** Illustration of a simple HMD projector based on collimating a micro-display panel.

The image, located at infinity has height (exit pupil) of the projector is determined by the diameter of the lens, whereas the field-of-view  $\Theta$  within the projector is determined by:

$$\tan(\theta/2) = x/2f$$

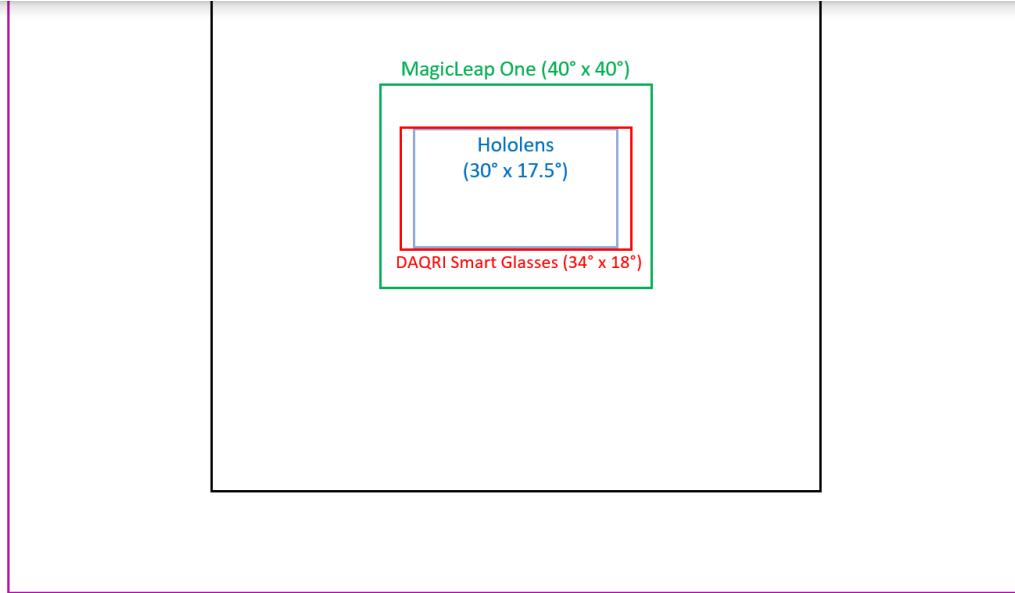
To increase FOV for a given display, we need to decrease  $f$  or increase  $x$  however etendue tells us increasing the solid angle reduces image size. In addition, the semi-diameter of the lens cannot be greater than its radius of curvature, which is what determines its focal length (it is common to use a mirror, so there is only one surface



because they allow “pupil expansion” or “pupil replication” such that the etendue relation is not directly impacted - however this has other consequences, such as on image quality, efficiency and luminance, as previously mentioned. Additional complications also arise when using scanning laser systems for projectors, since the exit pupil of the projector is very small. One method of expanding such a projector is to use an intermediate screen that then acts as a secondary source, this however adds bulk - additional relay lenses are required, adds speckle and also reduces efficiency. Another pupil expanding method is the use of waveguides, however artifacts are very difficult to suppress without eye-tracking and active-correction.

### **Field of View (FOV)**

While everyone seemingly wants more FOV, this too has to be carefully balanced against other aspects. The need for more FOV strongly depends on the use cases the glasses are designed for. Consumer use cases (e.g. playing games) benefit significantly from more FOV due to increased immersion. Many professional use cases (maintenance, inspection) are just fine with a FOV of only  $40^\circ \times 30^\circ$ , since the focus area is small and other aspects (such as increased safety due to unobscured peripheral view) are more important.



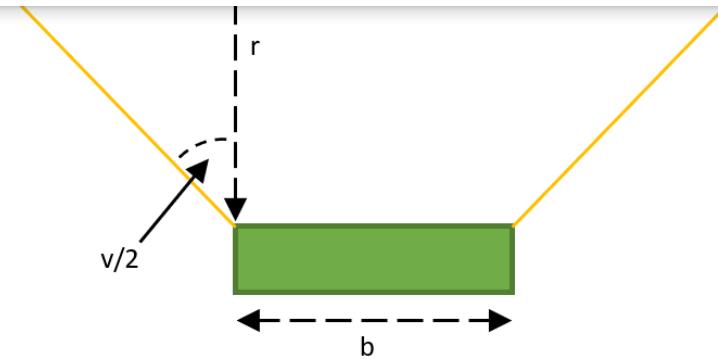
**Figure:** Field of view of various AR devices vs VR devices and human field of vision. For simplicity all FOVs are drawn as rectangles. Note that in reality neither VR devices nor the human field of vision is rectangular.

Still, as can be seen in the figure above, today's AR displays cover only a tiny fraction of the full human field of vision. VR devices, which are much bulkier and benefit from significantly simpler optical setups are much closer to covering the full human field of vision.

Field of view, eye box size and eye relief (see below) are closely related through the (simplified) formula

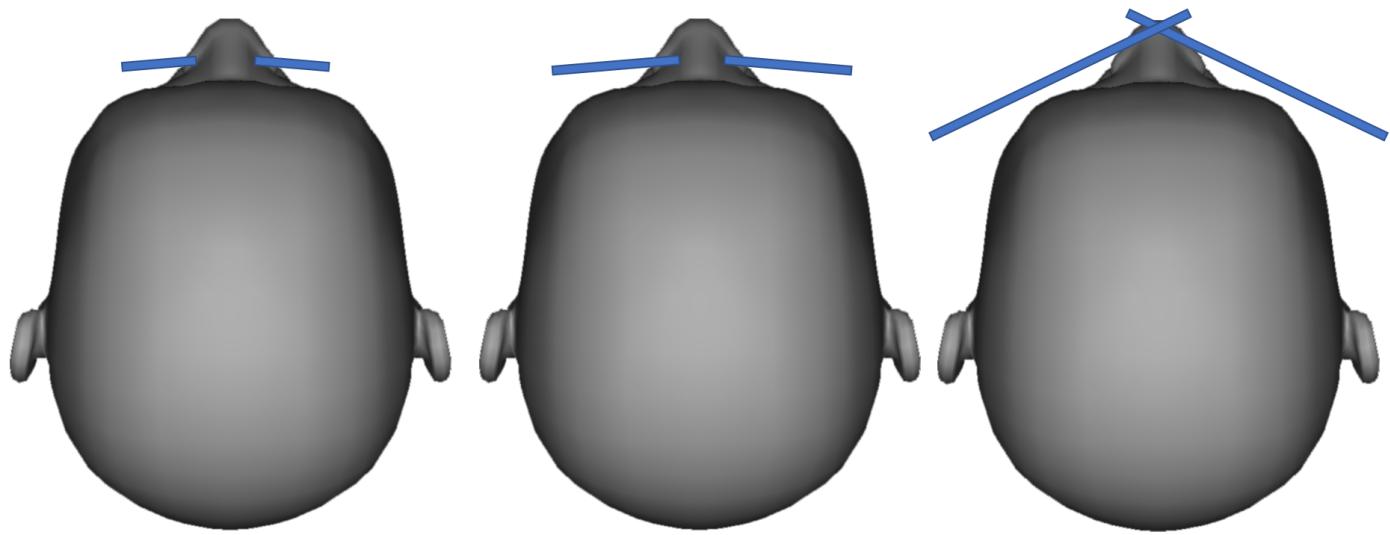
$$s = b + 2r \tan(v/2)$$

where  $s$  is the size (e.g. width) of a planar optical surface (e.g. a waveguide),  $b$  is the eye box size,  $r$  is eye relief and  $v$  is the field of view. The following figure visualizes this relation.



**Figure:** Relation between optical surface size ( $s$ ), eye relief ( $r$ ), eye box size ( $b$ ) and field of view ( $v$ ).

For an exemplary horizontal field of view of  $40^\circ$ , 20mm eye relief and 20mm eye box size, the display surface needs to be 35mm wide. For a desired display with  $90^\circ$  FOV (and otherwise same specifications), the display already needs to be 60mm in size. For a waveguide with  $150^\circ$  (diagonally) as announced by DigiLens, the display would need to be 170mm in diagonal. With a 4:3 form factor that display would be roughly 135mm x 100mm in size - per eye!



**Figure:** Waveguides sized 35mm (left), 60mm (middle) and 135mm (right)

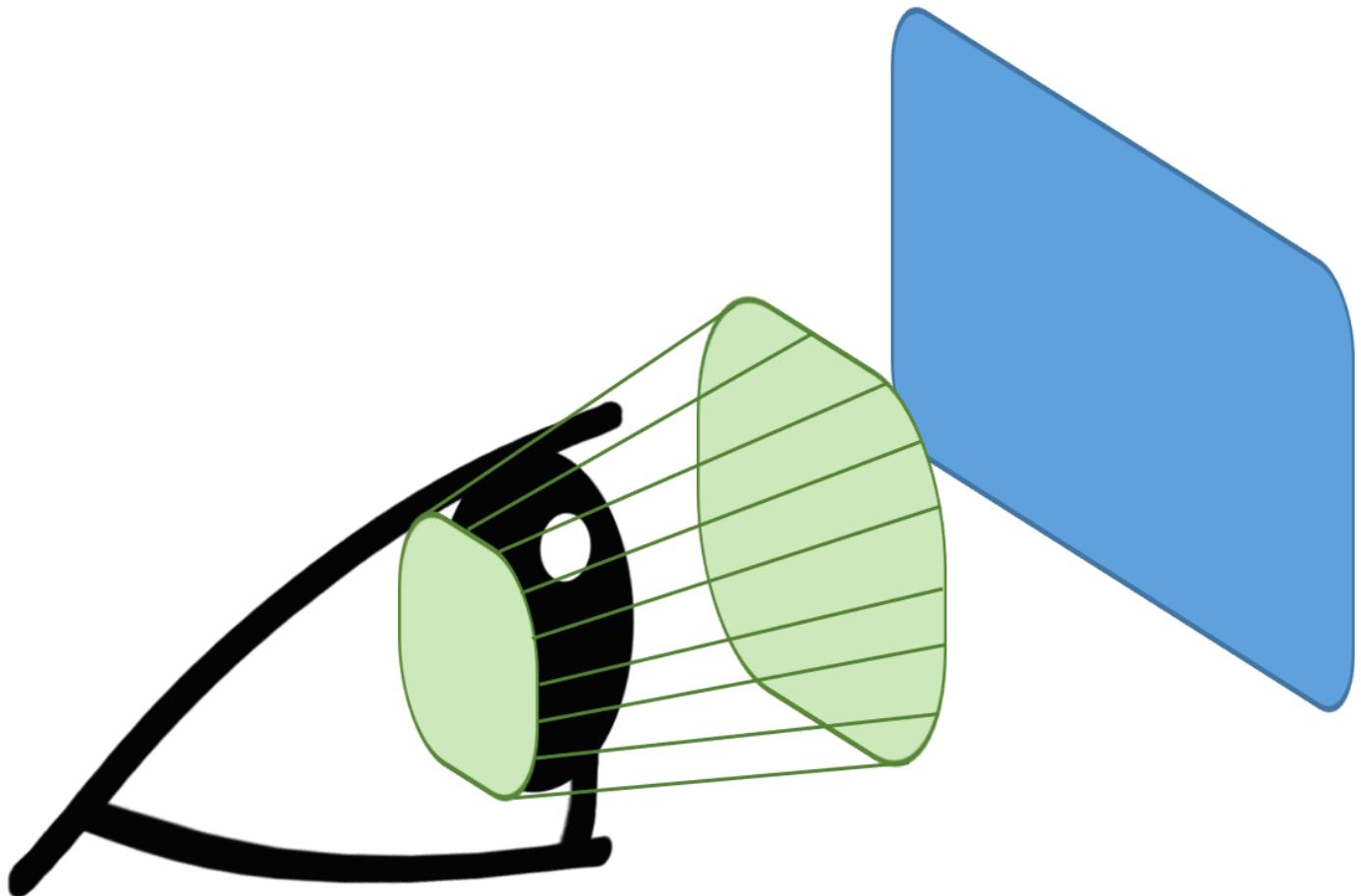
As can be seen in the figure above, even a 60mm wide waveguide is already quite large, but the 135mm waveguide simply does not work anymore. Not only would it be ridiculously large, but such a wide FOV must include the binocular overlap (see



and far from practical.

### Eye box size

Eye box is the area in front of a near eye display where the display content can be observed “properly”. Outside the eye box the display content might be distorted, colors might be wrong or flipped - or the content might not be visible at all. Despite being called “eye box”, the 3D shape of this area is not a box but rather a conic, becoming thinner with larger distance to the display. While eye box is a term often used for AR and VR displays, the more commonly used term in optics is “exit pupil”, where the instantaneous FOV is equal to the total FOV.



**Figure:** Eye (black) inside the eye box (green) in front of a display (blue)

Most people have already observed the issue of a too small exit pupil when looking through binoculars or a microscope: Unless each eye is positioned exactly at the right



So how large does the eye box have to be? The minimum practical size of the eye box is obviously the size of the pupil of the human eye (usually assumed to be ~4mm) - which is often the design goal for binoculars as mentioned above. Binoculars have a very small FOV so the pupil will stay relatively fixed. However, in the case of AR/VR displays the user's eyes will move depending on where the user focuses on the display. To support such eye motion the eye box size needs to be increased by at least a few millimeters in each direction.

Eye motion is not the only requirement for an enlarged eye box. The interpupillary distance (IPD) varies among the human population. To support this variation either mechanical adjustments (as in the case of binoculars) or optical adjustments (by further increasing the eye box width) are required. Moving parts are highly problematic in mobile devices, especially when accurate calibration is required. Since mechanical adjustment is not a good option, the eye box width needs to be further increased by at least 10mm, ideally rather 20mm.

Since the input pupil of a waveguide is usually not that large various techniques have been developed to increase the size of the exit pupil, such as pupil expansion or pupil replication. Diffractive waveguides have a natural advantage here, whereas with reflective waveguides expanding the eye box in both directions (2D pupil expansion) is quite challenging. As a result, semi-reflective waveguides usually have smaller eye box sizes than diffractive ones.

Designing a large eye box is generally challenging and has strong implications on other design parameters. E.g. the size of the eye box directly adds to the required size of the display area (waveguide). Also, as pointed out in the example with binoculars, a larger eye box requires more light output in order to achieve the same perceived brightness, which leads us to the next topic.

### **Brightness, Transparency and Duty time**

Display brightness - simply speaking - defines whether the display is bright enough such that the virtual content can be clearly seen in a certain situation. Transparency



topics are largely independent from a user's point of view, but technically so closely related that we'll discuss them together in this section.

**Display brightness** is certainly a highly challenging topic for mobile AR devices today; almost all AR HMDs today are severely limited in their display brightness: Hololens and DAQRI Smart Glasses have a display brightness of around 300nits, whereas the Magic Leap One is at around 200nits only. Because display brightness is so challenging, most AR glasses today are usually tuned to be barely bright enough for indoor usage and quickly become unusable outdoors - especially in direct sunlight. To ease things most AR HMDs use tinted visors that reduce **transparency** - and hence the amount of environment light reaching the user's eye - making the display relatively brighter. While this might be acceptable for consumers, in many professional areas low display transparency is not tolerated.

On top of tinted visors most optical designs block significant amounts of real world light. E.g. a birdbath design (as formerly used by ODG and now nreal) naturally blocks most of the incoming light.

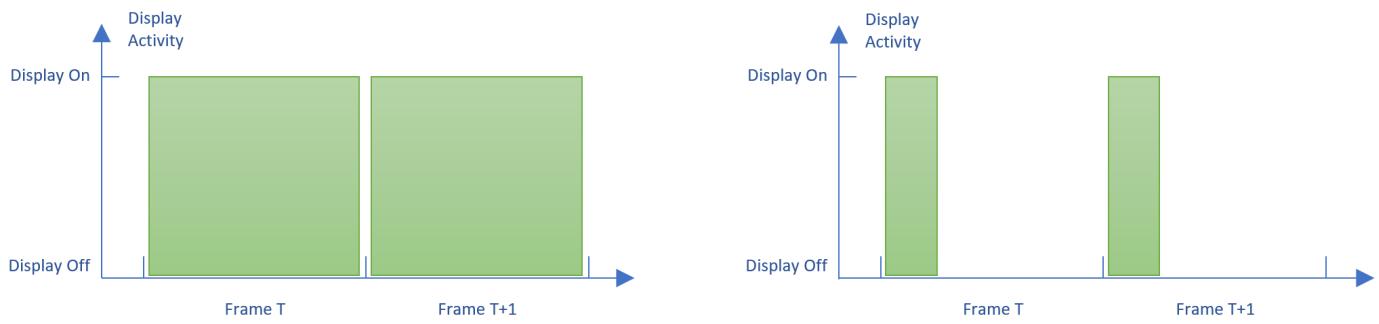
Karl Guttag's blog regularly reports on this aspect ([here](#) and [here](#)). According to Mr Guttag Hololens allows only ~40% of the environment light to reach the eye. In the case of the MagicLeap One it is only ~15% and in the case of the nreal glasses presented in early 2019 it is only around 25%. Fortunately though, the human eye has an enormous dynamic range of around 1:109 and can therefore deal quite well with such brightness reductions. Still, in dark environments these brightness reductions can be problematic.

Since **duty time** defines how long a pixel is allowed to emit light per frame it is closely related to display brightness. A duty time of 100% would mean that the display emits light continuously; for a 60Hz display that would mean around 16ms per frame. However, showing the same pixel values for such a long time would result in motion artefacts ("smearing"). E.g. let's assume a display with 40° horizontal FOV, 1280 pixels resolution along the horizontal direction and a head rotation of 60° per second. In this scenario the head rotates by around 2 pixels per millisecond. With a



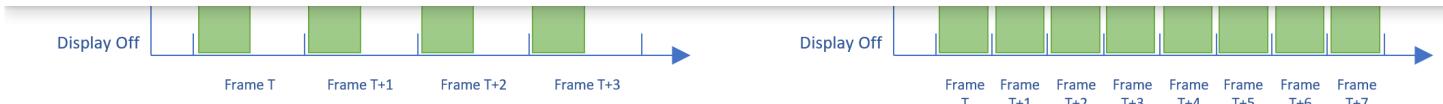
**Figure:** Same image without (left) and with (right) horizontal motion blur

Fortunately though, when looking at details humans hold their head much more steady, so in practice a pixel duty time of 4ms or less is still usually acceptable for AR displays.



**Figure:** Duty time for 60Hz frame rate; Left: 16ms duty time; Right: 4ms duty time

LCOS panels are lit with separate LEDs, which can be very bright, thereby enabling bright displays with high transparency and duty times of less than 1ms (per color). OLED displays on the other hand are much dimmer and hence quickly run into severe problems when duty times are short and transparency is required be high simultaneously. The figure above shows example duty times for an assumed 60Hz frame rate. The area of the green bars directly reflects display brightness. In the left image with 16ms duty time the display is bright, but suffers from smearing. In the right image with 4ms duty time smearing would be much reduced, but the display would potentially be too dark.



**Figure:** Duty time of 4ms for 120Hz (left) and 240Hz (right) frame rate

At a frame rate of 60Hz (~16ms), an absolute duty time of 4ms would result in having the display pixels only active for ~25% of the time. As a result, the display brightness would be reduced by a factor of four. If the display was running at 120Hz instead (see the left image in the figure above), 4ms duty time would mean pixels still emitting light at ~50% of the time. Similarly running at 240Hz (see right image in the figure above) would enable having ~100% duty time and still meet those 4ms. However, such a high frame rate is unrealistic in the case of mobile devices same as displays panels actually supporting it.

## Contrast

There is no official definition or way to measure contrast or contrast ratio. Still most people have a sense for what contrast means. Simply speaking contrast describes a display's ability to produce brighter and darker pixels simultaneously and can e.g. be defined as

$$\text{contrast} = \frac{\text{brightest} - \text{darkest}}{\text{brightest} + \text{darkest}}$$

If a display has poor contrast it will not be able to show bright content as well as dark (transparent) content simultaneously. Depending on the overall brightness of the display, it will therefore either not be able to show bright areas or supposedly dark (transparent) areas will not be dark. In optical see-through AR displays the role of black is replaced with transparency: In AR displays with poor contrast one can therefore observe transparent areas showing up as dark gray instead. Putting it differently: As long as AR displays are not bright enough yet, the negative effect of



have a low contrast to start with - usually 1:100 to 1:200. QLEDs have a comparatively high contrast of 1:1.000.000 or more, which is why they are popular for home TVs nowadays. However, in both cases optical elements (prisms, lenses, waveguides) will lead to further contrast reduction so that e.g. the final contrast of an LCOS based system can easily be way below 1:100.

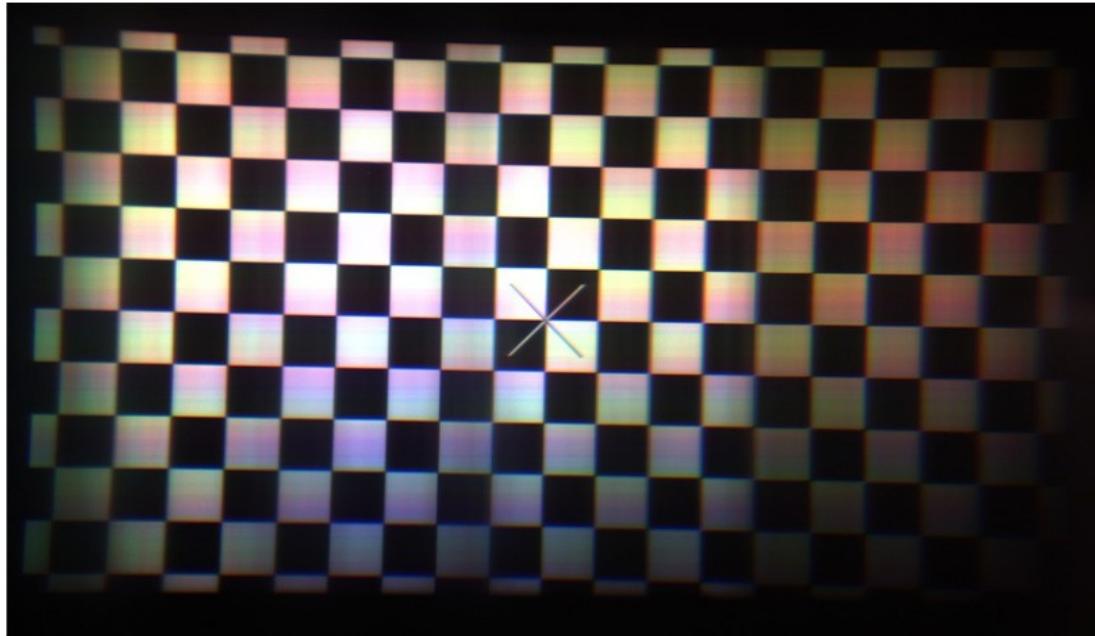
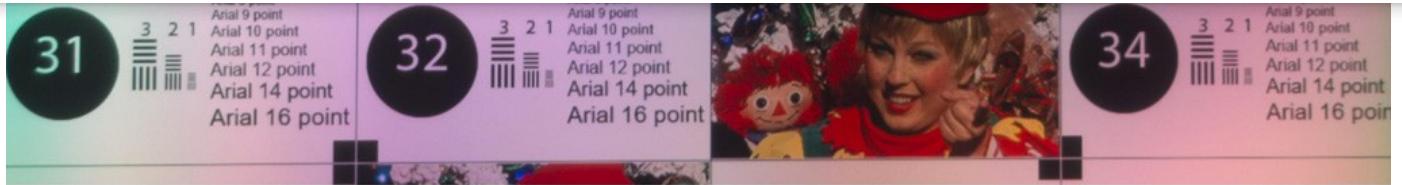
### Uniformity and Color Quality

**Color quality** defines how accurately a display is able to reproduce colors - e.g. does a pixel rendered as red by the GPU really look red on the display? To achieve proper color reproduction calibration (including gamma) is required. Since AR displays are typically of additive nature, the perceived color depends also on the scene on top of which the virtual content is overlaid.

Same as contrast, color quality can vary significantly depending on the location on the display: e.g. a pixel color can appear quite different depending on whether it is more on the left or on the right side on the display. These artefacts are often view dependent, meaning that the position of the user's pupil also has an influence and would require eye tracking to resolve.

**Uniformity** describes how much pixel colors vary depending on their location on the display: On a perfect display every pixel rendered with the same RGB value would look identical. In practice though brightness, contrast, color and other properties vary depending on the location on the display as well as the angle under which pixels are observed.

While uniformity is often pretty good on freeform combiners (such as in the Meta2) as well as semi-reflective waveguides (such as from Lumus), diffractive waveguide displays (such as in Hololens and MagicLeap) suffer noticeably from uniformity issues. Such displays filled with mid-level gray pixels show all kinds of color tones (see the figure below).



**Figure:** Top: Color issues on a diffractive waveguide (image courtesy of [Karl Guttag](#)); Instead of a uniform white or grey tone, the display shows various hues. Bottom: Significant brightness non-uniformity

## Resolution

Display resolution describes how many distinct pixels can be displayed. It is one of those properties a lot is written about, but understanding is comparably limited. The ultimate goal for display resolution is reaching or going slightly beyond the human vision limit of roughly one arcmin ( $1/60^\circ$ ).

Due to a drive to push specs more and more, many phones today have much higher resolution than the human eye can observe under normal conditions. E.g. looking at a phone with a display size of 14cm from a distance of 40cm means that the phone's screen is  $\sim 20^\circ$  in the person's field of view, hence not requiring more than 1200 pixels along the longer side. Still many phones today have 50% more display resolution than this.



billion pixels per second...

On mobile phones users look directly at the screen. So (ignoring protective shields) there are no optical elements that would affect the pixel quality delivered by the display panel in a negative way. In AR and VR devices though a complex optical system sits between the user's eye and the display panels that can seriously degrade image quality. The perceived resolution (what reaches the eye) can be significantly lower than the display panel's resolution. E.g. as pointed out by **Karl Guttag**, the MagicLeap One's effective resolution is only about half of its panel's resolution (VGA instead of HD), similarly the Hololens display loses a lot of resolution along the optical path from LCOS to eye.

Hence, as long as the optical system - e.g. the waveguide - is the limiting factor quoting display panel resolutions is usually not meaningful.

### Real-world Distortions

In the case of optical see-through displays the real world is observed through the optical elements of the display. In most AR devices those elements make up a subset of the following

- The waveguide showing the virtual image required for augmentations (e.g. Hololens)
- Or a freeform combiner reflecting a projected image into the eyes (e.g. Meta2)
- A visor on the outside of the device protecting the interior electronics and optical elements (all waveguide based devices)
- Push/pull lenses that move the virtual image focus plane from infinity to something more practical such as 2m (reflective waveguides e.g. in DAQRI Smart Glasses)
- Additional plastics that protect the fragile waveguides on one or both sides (e.g. Hololens)

Some of these elements have additional, undesired optical properties. E.g. a waveguide is built to bend (guide) light into the right direction. However, as such it



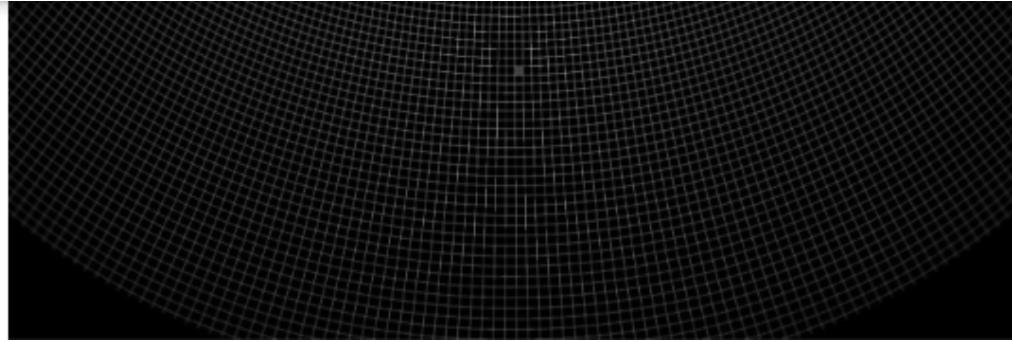
Naturally, one would want as little distortions of the real world view as possible, but in practice the desire to limit weight and cost requires compromises that can generate noticeable artifacts.

### **Virtual Image Distortions**

Optical engineers aim at designing optical paths with maximal possible image quality, which also includes minimizing distortions: In a perfect case the rectangular pixel grid of the display panel would appear as an equally rectangular pixel grid towards the user. In a direct-view scenario such as looking on a regular mobile phone this is trivially given. In a complex optical setup such as in AR displays image distortions will often have to be tolerated in order to further optimize other parameters.

Fortunately, compared to real-world distortions, distortions to the virtual content can be handled more effectively as long as the display is well calibrated (and the distortions are largely view-independent). With proper calibration these distortions can be handled as part of the rendering pipeline (digital undistortion) with low or no additional processing cost. Still, depending on the amount of optical distortion in the system this can lead to noticeable artefacts such as reduction of the display resolution in certain areas.

Generally, distortions of waveguides are usually quite low such that at least for consumer use cases they might even be ignored (not calibrated and not corrected digitally). Free-space combiners - same as VR displays - usually generate severe distortions that need proper handling. Since the distorted image can be strongly non-rectangular in this case the effective area of the display panel might be reduced as well. The figure below shows an exemplary distortion grid of a free-space design. As can be noticed part of the grid falls outside the display panel and part of the display panel cannot be observed by the user (black area without grid). The image also shows well the resolution difference between top and bottom.



**Figure:** Exemplary distortion of a free-space combiner. This picture was generated by tracing the image of a rectangular grid along the optical path onto the display panel.

### Eye Safety

Two types of eye safety are important when talking about AR displays: keeping the eyes safe from the AR display and using the AR display to keep the eyes safe from external sources of harm.

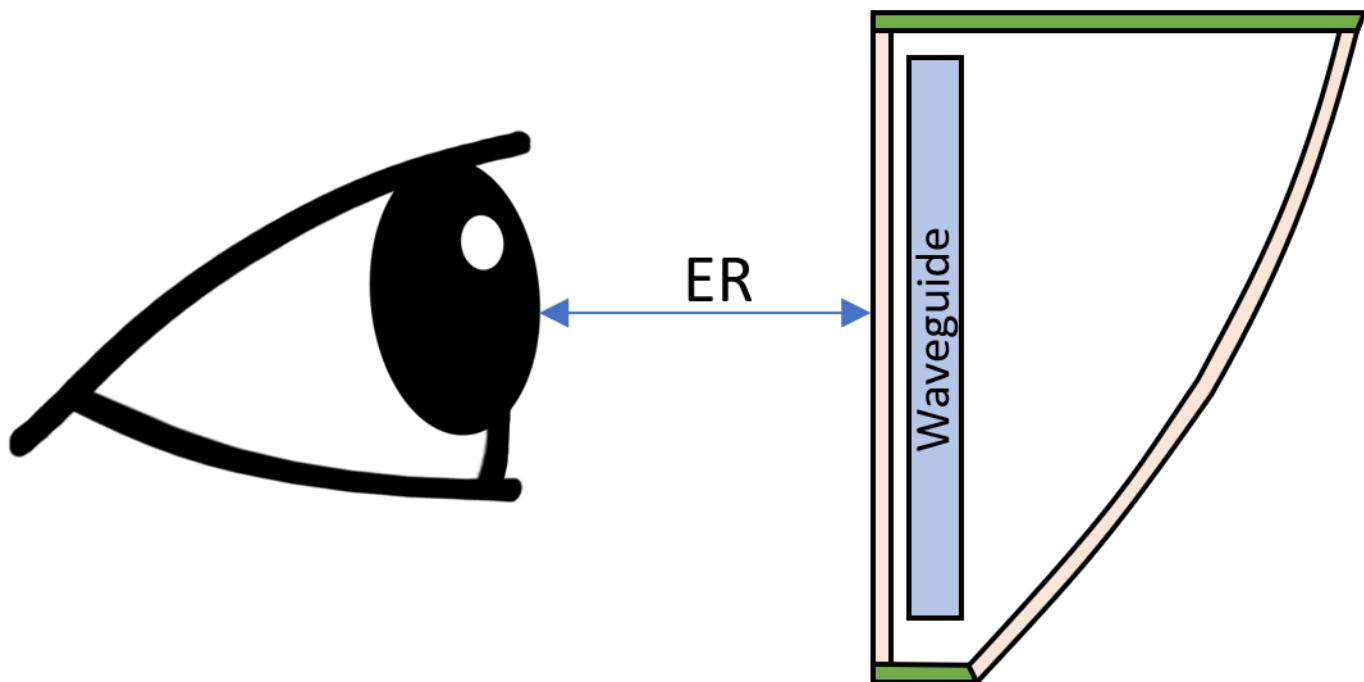
Keeping the eye safe from the AR display sounds like a no-brainer. Any product may it be consumer or professional level obviously has to fulfill such requirements. However, in the case of near-eye displays that are placed just a few centimeters away from one of the most vulnerable human organs, special care is mandatory. This becomes even more important as many AR displays use glass elements as part of their optical stack. Upon impact these glass elements can break and hurt the user. Hence, all glass elements need to be put into a protective cover that is unlikely to break in most conditions.

While this might sound obvious, it is not always the case. E.g. the recently announced Lenovo ThinkReality glasses place their reflective waveguides directly in front of the user's eye without such a cover. Since these waveguides are built from many small glass elements glued together in horizontal stripes they could easily break and harm the wearer.

Keeping the eye safe from external forces is a requirement usually only dealt with in commercial and industrial environments. Here eye protection safety standards such



Same as for eye box, there isn't a commonly agreed upon definition for eye relief. Simply speaking it is the supported distance of the pupil to the closest point on the AR display. Since not all users have the same head shape, a certain range of eye relief needs to be supported in practice, thereby defining the thickness (along the viewing direction) of the eye box.



**Figure:** Eye Relief (ER) is the distance between the pupil and the closest optical surface.

In general an eye relief large enough to wear regular prescription glasses is preferable so that users who require eyesight corrections don't need to buy lens inserts for their AR glasses. However, as mentioned earlier, the eye box is actually not a box, but conic shaped instead and becomes thinner with larger distance to the display. Hence, supporting a large eye relief and an eye box with sufficiently large width and height is challenging.

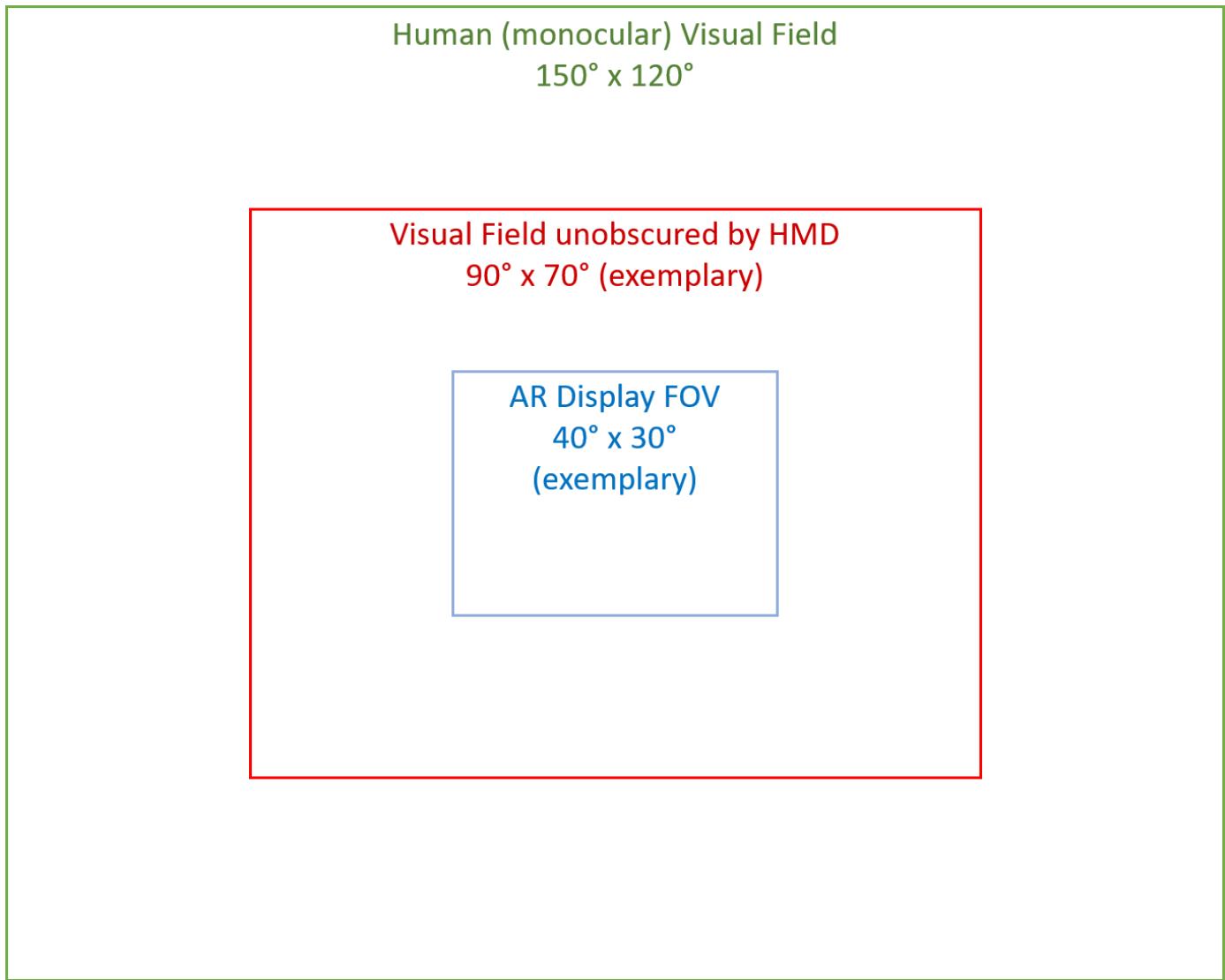
### Peripheral Vision

When it comes to AR glasses, not one but two fields of view are important: The FOV of the augmentable area is the part of human visual field where the glasses can show



The natural human visual field is around  $150^\circ \times 120^\circ$  per eye and  $220^\circ \times 120^\circ$  with both eyes combined.

Mounting a display in front of the eyes will naturally lead to additional obstruction, so an important design goal is keeping this obstruction to a minimum. The figure below shows how the natural view (green), the unobscured view (red) and the augmentable view (blue) roughly relate in size on today's devices. For simplicity all the areas are drawn as rectangles.

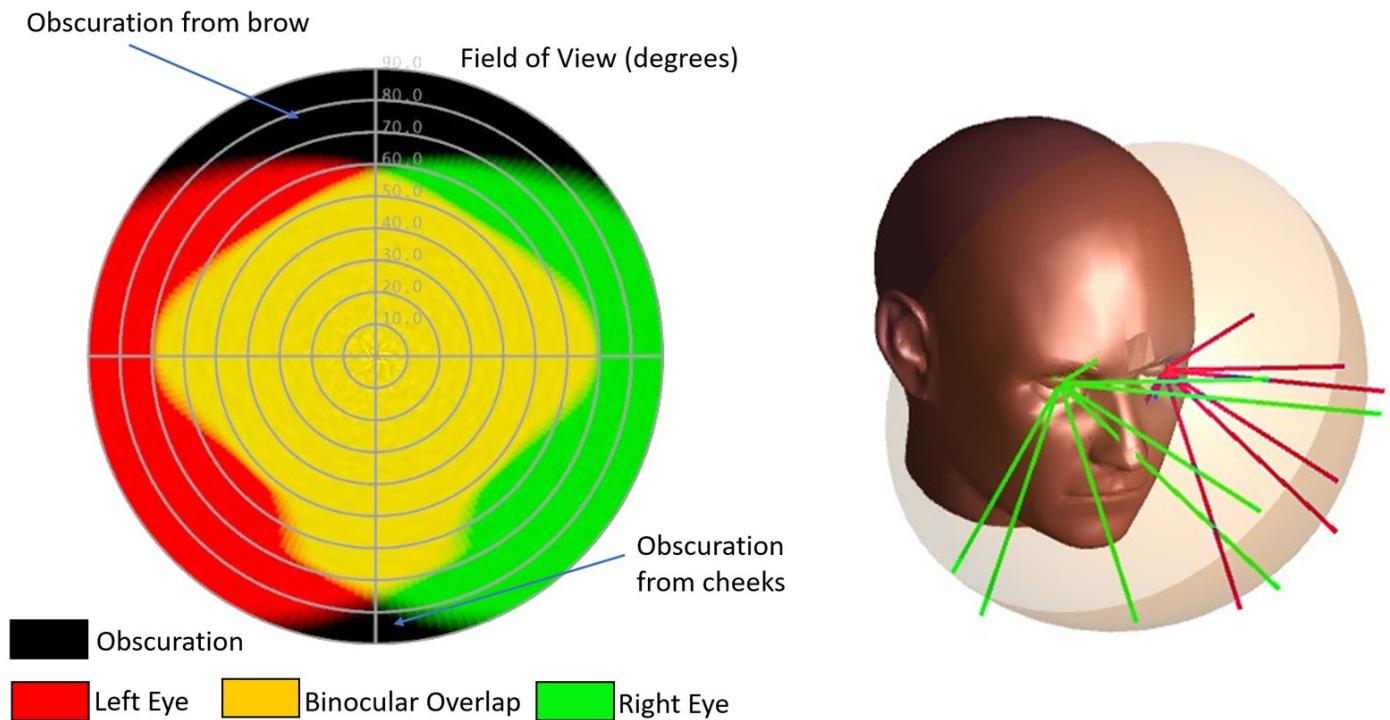


**Figure:** Comparing the human visual field (green) against an exemplary visual field of an AR device and the actual augmented view. The area between green and red represents the visual field that is blocked by the device's frame. The area between



goal is to maximize the non-obscured field of view (red area above). To achieve that, anything that can block the view needs to be pulled outwards. This includes parts of the display (such as the projector) as well as other elements such as sensors or supporting structures such as the glasses' arms.

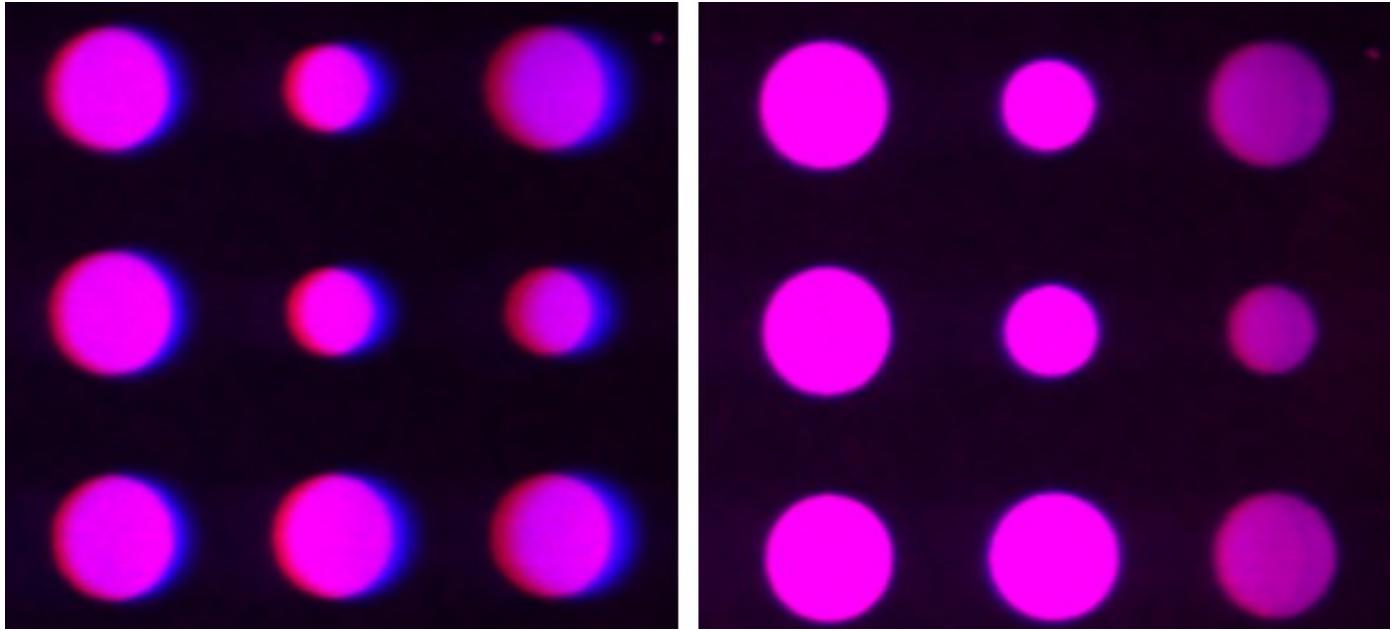
Different to our simplified visualizations above, the visual field is not rectangular. As the figure below shows the visual field is mainly limited by eyebrows, nose and cheeks: The combined red and yellow area depicts the left eye's visual field. Similarly, the combined green and yellow area depicts the right eye's visual field. The yellow area depicts the binocular overlap - the field that both eyes can observe.



**Figure:** Human visual field for left and right eyes (left image). The left image was generated doing raycasting using a virtual head model (right image)

### Chromatic Aberrations

The refractive index of a lens varies with the wavelength of light, which results in different “color-dependent” focal lengths. In cameras this is usually compensated for by combining multiple lenses, but due to the size constraints this is often not possible in AR displays. Hence, chromatic aberrations are a noticeable concern in AR



**Figure:** Left: Red and blue breaking up due to chromatic aberrations. Right: The same image digitally corrected by warping each color channel accordingly.

## Depth Perception

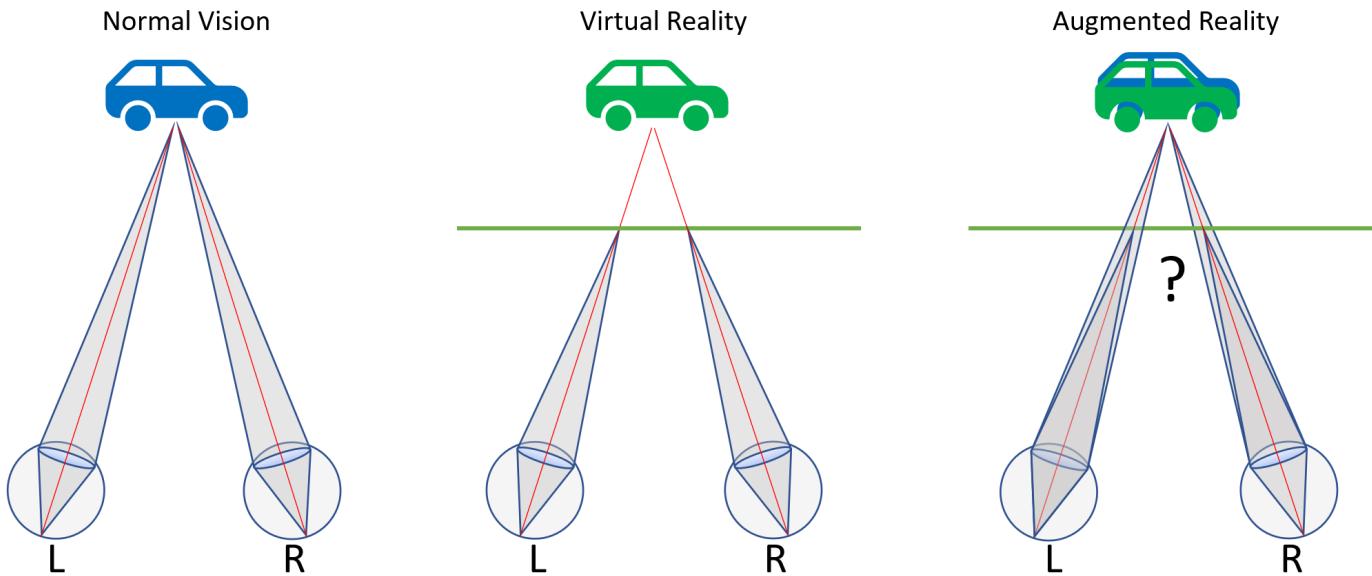
There are multiple cues in human vision that allow us to perceive depth. For AR displays the two most important cues are vergence (our eyes rotating to look at the same object) and accommodation (our pupils focusing at an object), which are neurally coupled. Vergence and accommodation not matching results in discomfort, called the vergence and accommodation conflict (VAC).

Most people will have noticed VAC when watching a 3D movie: while focus never changes (the TV or projection screen doesn't move...) one still experiences a 3D effect due to our eyes seeing slightly different images (stereoscoping content). In a cinema the focus plane is given by the room setup: If a person sits 10 meters away from the projection wall then the focus plane is fixed at 10m. At this distance humans can barely distinguish distance anymore based on pupil focus. Hence, as long as the stereoscopic content also resides at this distance or more (rather than popping out) things look natural.



The figure below highlights the differences for normal viewing, Virtual Reality and Augmented Reality: In the case of normal viewing vergence and accommodation are in sync - both adjust to the same distance. In the case of Virtual Reality accommodation is always at the same distance (usually around two meters) whereas vergence depends on the screen content rendered in stereo. In the case of Augmented Reality the conflict can be even larger: An object augmented with virtual content will appear in sync with respect to vergence, but accommodation for the real and virtual object can be very different.

Ideally we'd be able to select a different focus distance per pixel and experimental systems going into this direction have been demonstrated. However, it will take long until such systems reach commercial level.



**Figure:** Vergence and Accommodation in normal viewing conditions (left), Virtual Reality (middle) and Augmented Reality (right).

As long as we have to live with a single focus plane AR display designers need to decide where to place it. The best fit for most scenarios seems to be somewhere at around 2m. This focus plane should be roughly flat and identical for all colors. This is not a trivial design goal though and hence, when measuring today's AR displays one



Display size and generally glasses size is one of the most challenging design parameters of AR glasses today. Since both FOV and eye box shall be large it is very hard to make the display small. It is like asking for a truck that is small, but has a high capacity for transporting goods at the same time. Large displays lead to bulkiness, which leads to less practical glasses; the larger the glasses are the more likely the user is to bump into something.

However, larger displays usually also result in heavier optics. Due to quality and refraction index requirements many optical elements of today's AR displays are made of glass, which can quickly become too heavy as size increases.

In his speech at Oculus Connect 5 in 2018, Michael Abrash postulated that AR glasses must not weigh more than 70 grams. Size and weight are not parameters independent of other properties such as form-factor though. The human head can comfortably carry significantly more weight than those 70g - if that weight is well distributed. Whereas the nose bridge quickly hurts already when carrying very little weight, ears can already carry more weight and the head top even more. Weight distribution is more important than weight. E.g. the Meta2 glasses, while not heavy absolutely, put a lot of pressure onto the forehead due to their unfortunate weight distribution.

## Optical Efficiency

Optical efficiency is about how much light sent out by the light emitting elements (e.g. LEDs) actually reaches the user's eyes. It will be surprising to most that today's waveguide-based displays are extremely light inefficient, with most of them having a light efficiency of only about one percent. Fortunately projectors combining LCOS and LEDs are bright enough to provide sufficient amount of light for waveguides. OLEDs however are therefore not an option for waveguides today. With combiner displays (like Meta2) on the other hand, the light efficiency can be controlled well by the amount of transparency the combiner element has: The more reflective the combiner, the more light efficient it will be. This will, however, also result in more environment light being reflected and hence less environment light reaching the eye (reduced transparency).



brightness levels. Micro LED panels with sufficient resolution and panel sizes suitable for AR displays have recently been demonstrated, but so far these panels are monochromatic only. It will probably still take several years to get similarly spec'd panels with full RGB support.

## Latency

Motion to photon latency defines how long it takes from an event (a motion) until the display shows a respective update. E.g. when the user rotates the head to the right then the content on the display must "shift" to the left accordingly. Latency is not a well studied topic in AR, largely because systems with a low enough latency have not been available until recently. However, it is commonly agreed upon that a latency of 5 milliseconds or less is sufficient for optical see-through displays.



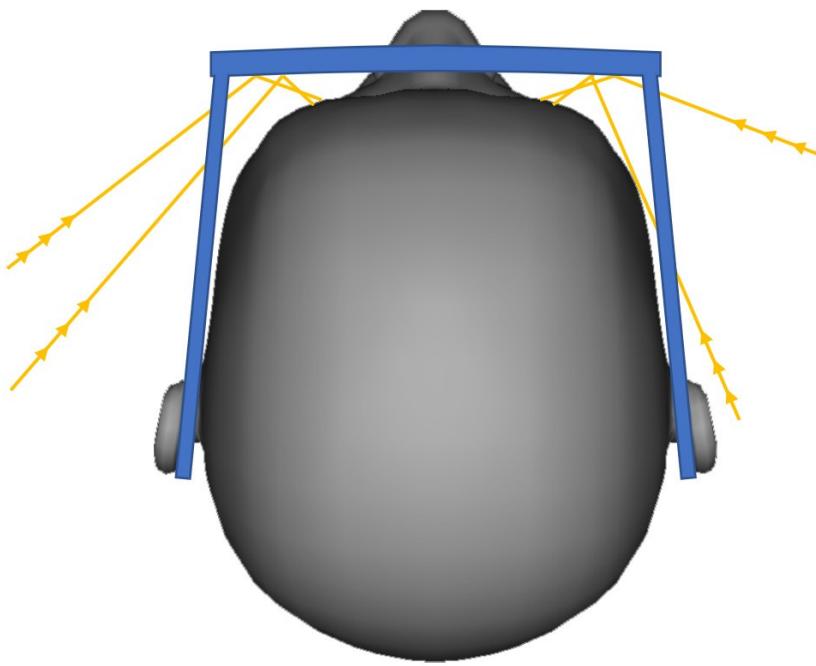
**Figure:** Due to latency labels can shift from clear locations (left) to unclear or wrong locations (right) during fast head motions.

Besides algorithms and other electronics, latency is mainly a function of the display panel (OLED, LCOS) and the display protocol (MIPI, DisplayPort, HDMI). Latency and the choice of display panel is a complex topic as it has huge effects on the electronic and software design. E.g. OLEDs, which do line-sequential ("rolling") updates, require completely different strategies for data transfer and motion compensation than LCOS, which do color-sequential ("global") updates.



Most users and glasses makers today dream of AR glasses with a form factor similar to that of sunglasses. While this sounds like a no-brainer there is an important issue that is usually overlooked: stray light.

The more open AR glasses are the more light will enter the system from unwanted directions and sources. While AR displays are usually coping well with environment light coming from the front, light coming from the side or from behind the user causes severe issues. Regular prescription glasses generally deal well with this by not reflecting much light in general. AR displays however have to reflect and bend light in order to function and thereby become significantly more affected by stray light. Diffractive waveguides suffer specifically from this, showing light coming from the side in the form of rainbow artefacts on the display (see right image in the figure below). Reflective waveguides perform better, but are also not free from issues. In some designs stray light can be reduced, but in others the problem is not easily resolved as in the case of diffraction techniques.

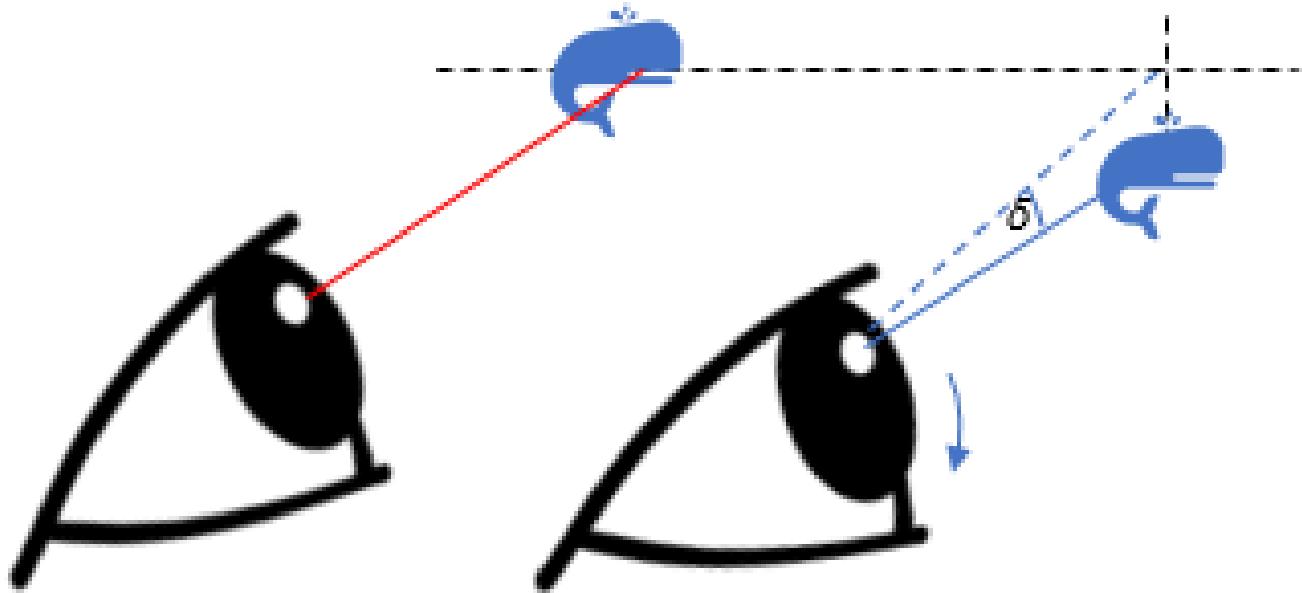


**Figure:** Left: Stray light from the side reflecting into the user's eyes. Right: Artefacts due to stray light from a ceiling lamp.

## Visual Comfort

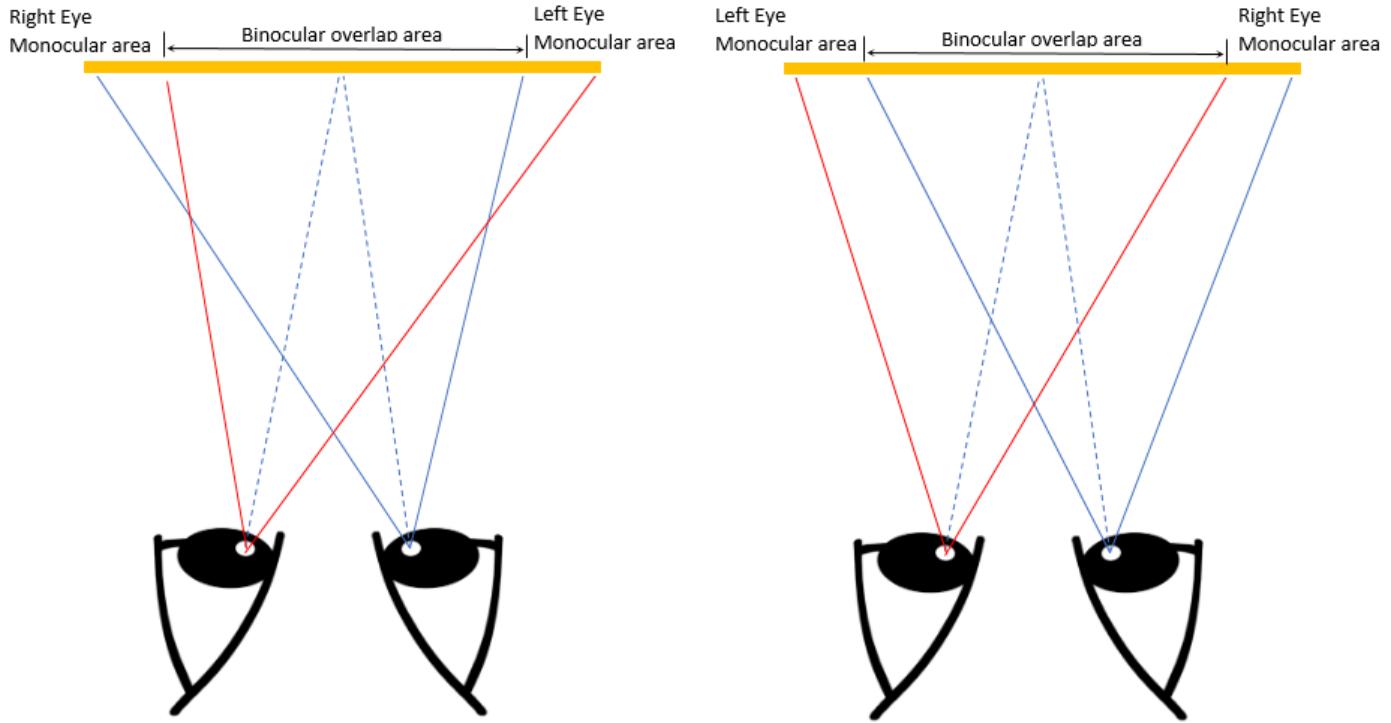


significant impact on comfort. One of these is known as **dipvergence**. Dipvergence arises when there is a vertical disparity or tilt between a binocular pair of displays. The human vision system is known to be intolerant to this, and can result in dizziness, nausea and even vomiting. Sometimes on the field people assume VAC is the cause of such symptoms when in fact it is slight misalignment between displays. A study by Self in 1986 in reference to US Navy training texts, points out that the vertical misalignment  $\delta$  of the axes of binocular barrels should not exceed 2 arcmin to avoid eyestrain.



**Figure:** The vertical misalignment of images, dipvergence, which can result in viewing discomfort. The right eye needs to rotate slightly in order to fuse the binocular image.

Another area that can impact comfort is the amount of **binocular overlap**. It is not necessary, for example to completely overlap left and right image fields. In fact, it is common to increase the effective FOV by deliberately not overlapping the display, in one of two ways: Divergent and Convergent overlap. Generally the human visual system can tolerate this, since the image of the real world seen by the left and right



**Figure:** Partial overlap schemes, Left: Divergent partial overlap, Right, Convergent partial overlap.

**Centre-of-mass** is another important element, although non-visual, which can cause a user unnecessary levels of discomfort in the neck through inappropriate design. The placement of display component and driver electronics needs to be arranged to help minimise any shift in the centre of gravity. If the user needs to look up and down over a range of angles, then this can be critical.

### What Else?

Beyond the topics covered in the sections above there are other design options that haven't even made it into commercial systems yet. Most devices today are capable of showing only a single focus plane at a fixed distance. The Magic Leap One goes one step further with two focus planes, but pays a heavy price in terms of reduced display quality and transparency. However, humans are able to distinguish around one dozen focus distances whereas the current approaches don't scale well beyond



block pixels in optical see-through displays. Today's passive optics are not able to do this since their real-virtual mixing works purely additive. For black pixels one would need to enable blocking the real-world light at a per-pixel and per-frame level. While an LCD layer might come into mind, that approach would reduce the display's transparency by half, create issues with polarized environment light and is hence usually not a viable solution.

We also haven't covered the topic of power consumption and heat dissipation. Humans are highly sensitive to heat sources close to their eyes and the face in general. Hence, a headworn device should not dissipate more than one Watt in the face and temple areas. Basically all AR devices today struggle already with generating too much heat. So in order to make displays noticeably brighter the optical designs will have to become significantly more efficient rather than just cranking up the power of the display's light sources.

### What To Expect from the Future

The common ask for making the field of view larger as well as making the glasses generally smaller (towards sunglasses form factor) clearly works against improving the many parameters we discussed above. Different to electronics, miniaturization is usually not an option or benefit in optical design, since it will result in shrinking parameters such as focal length, eye box size or eye relief as well.

It is unlikely that new break-through technologies will push forward more than just a small number of the design parameters discussed above. E.g. the advent of full color microLEDs will enable smaller and brighter projectors, but this has only a limited impact on the overall display size. Techniques in the field of optics on the other hand are rather unlikely to change drastically.

It will therefore take quite a while until AR glasses with a sunglasses formfactor, large field of view, high brightness, outdoor suitability and all those other aspects visionaires dream about will come to life. Instead, - same as in the field of battery technology - we will more likely see incremental improvements year over year.

1 503 · 59 Comments

[Like](#) [Reply](#)**Tanveer Singh**

1y ...

Best article I read on Ar after a long time. Thanks for boiling down complex concepts into basics.

[Like](#) [Reply](#) | 1 Like**Eric Snyder**

1y ...

Really, really good article to explain AR challenges, Daniel.

[Like](#) [Reply](#)**Tejaswini Shenoy**

1y

This is the best and most comprehensive article over XR displays that I have come across! Thank you, Daniel Wagner!

[Like](#) [Reply](#) | 1 Like**Esaias Pech**

1mo

Great explanation Daniel Wagner!

[Like](#) [Reply](#)**Richard Bailey**

2y

Great explanation Daniel Wagner!

[Like](#) [Reply](#)**Sagar Madliyavar**

1y

A very good description of different aspects in a single article. 🙌

[Like](#) [Reply](#)**Jay Quincey**

3y

This is one of the most accurate and relevant articles I've read on XR optics and their challenges to date. Awesome work!

[Like](#) [Reply](#) | 1 Like



徐凱

3y

<https://www.kura.tech/> Too amazing to believe such specs: 150°+4000nis+8K+95%. Is it true?

[Like](#) [Reply](#)

Minnie Ma

2y

徐凱 So do you think it is true?

[Like](#) [Reply](#)

徐凱

3y

<https://www.kura.tech/technology> from page bottom: "AR Display And Optics Technologies Comparison"

[Like](#) [Reply](#)[See more comments](#)

To view or add a comment, [sign in](#)

## More articles by this author

### An early Analysis of the Magic Leap One

Dec 21, 2017

### Others also viewed



#### The most important things to know about LCOS in 7 minutes

Florian Mayr · 2y



Embedded Systems · 1w

**What is assisted reality and why should we care about it**

Naveen Joshi · 2y

**★★★★★ #FPGA (Field Programmable Gate Arrays)**

Embedded Systems · 3w

**★★★★★ Printed Circuit Board #PCB Design: Fundamentals and Best Practices**

Embedded Systems · 3w

**★★★★★ Intelligent Irrigation Systems Based on #IoT**

Embedded Systems · 3w

**★★★★★ Indoor positioning**

Embedded Systems · 3w

**Apple's Secretive AR/VR Footprints Point to Something Big Next Year**

Balaji V. · 2y

**Three Startup Lessons From my Days at Early Auris**

Jeffery Alvarez · 7mo

## Explore topics

- Workplace**
- Job Search**
- Careers**
- Interviewing**
- Salary and Compensation**



© 2023

[\*\*About\*\*](#)

[\*\*Accessibility\*\*](#)

[\*\*User Agreement\*\*](#)

[\*\*Privacy Policy\*\*](#)

[\*\*Your California Privacy Choices\*\*](#)

[\*\*Cookie Policy\*\*](#)

[\*\*Copyright Policy\*\*](#)

[\*\*Brand Policy\*\*](#)

[\*\*Guest Controls\*\*](#)

[\*\*Community Guidelines\*\*](#)

[\*\*Language\*\*](#)