

Developing inexpensive and convincing stereoscopic-3D anatomy learning experiences: an artistic and technical reference

Rohit Malyala

April 22, 2019

1 Introduction

Stereopsis refers to the ability to perceive depth, stemming specifically from the disparity and integration of visual information from two laterally located eyes. Each eye receives a slightly different angle on objects in the field-of-view, causing *horizontal* or *binocular disparities*, which when integrated together, produce an understanding of environmental depth to supplement other visual factors such as motion parallax, contrast, overlap, and more. In anatomy education, spatial awareness and understanding of the distances and relationships between structures is paramount.

Most visual media is recorded by a single camera and projected onto a single screen (with no horizontal disparity). However, binocular disparities can be simulated in conventional video by showing two video feeds from laterally displaced cameras into each eye—virtual reality. So-called *stereograms* can be inexpensively both produced and delivered on smartphone devices, using VR solutions such as *Google Cardboard*. Hopefully, democratizing the production of high-quality stereograms delivered as accessibly as possible will improve understanding of spatial anatomy.

The following document characterizes methods and considerations, both artistic and technical, when creating a stereoscopic-3D anatomy learning experience. However, it is generally also suited to any type of close-subject stereoscopic media. All software used and referenced within is free and/or open-source, with the exception of some proprietary color-correction programs (which are fully optional).

Contents

1	Introduction	1
2	General Methods	2
2.1	Brief discussion on the learning benefits of VR stereoscopy	2
2.2	Creating a full video module, a step-by-step overview	2
3	Constructing a stereoscopic photoshoot environment	5
3.1	Dual smartphone camera apparatus	6
3.1.1	Toe-in vs. parallel camera setup	6
3.1.2	Focal length to distance and inter-camera distance	6
3.1.3	Hyper and hypo-stereopsis conditions	7
3.2	Adjusting for inter-pupillary distance	7
3.3	Resolution and framerate	7
3.4	Lighting, contrast, shadows	8
4	Labeling and artificial/digital stereoscopic 3D VFX	8
4.1	Artificial depth for text	8
4.2	Lines, moving towards or away from the screen	9

5	Shot selection and composition	10
5.1	Depth differences in a single frame	10
5.2	Depth difference between subsequent shots	10
5.3	Moving and panning shots	11
6	Color grading and post-processing	11
6.1	Equalizing color grade	11
6.2	Lateral borders	11
7	Publishing to the web	12
7.1	Differences between screen hardware	12
7.2	Publishing to YouTube (with metadata)	13
7.3	Conversion between SBS and color anaglyph	13

2 General Methods

2.1 Brief discussion on the learning benefits of VR stereoscopy

Virtual reality-based education for surgical residents, medical students, and anatomy students is on the rise. Purported benefits include improved immersion (e.g. CPR simulation games), and for anatomy, improved spatial understanding of objects from using stereoscopic learning experiences. Spatial understanding of anatomy is crucially important for clinical tasks, but is poorly understood. For example, identifying the gastroduodenal artery versus the duodenum itself is not necessarily clinically useful; however, understanding that they are spatially close to each other may lead to understanding that a duodenal ulcer may cause hemorrhage in the gastroduodenal artery. Unfortunately, not all anatomy faculties can produce or deliver physical plastic, plastinated, or cadaveric anatomy models to students.

Surgeons in particular display unusually good stereoacuity, and endoscopes are often already fitted with stereoscopic cameras for a true-stereoscopic video feed¹ (yes, even camera divergences of fractions of millimetres is enough to produce a stereoscopic effect!)². Hence, VR stereoscopic solutions, especially inexpensive solutions like *Google Cardboard*, are gaining traction.

High spatial reasoning ability of the student improves spatial anatomy comprehension³. Other factors, such as self-directed observation from multiple angles of an anatomical specimen, are also found to help improve learning. Again however, the anatomy learning benefits of stereoscopic VR, and the impacts of VR hyper or hypo-stereoscopy (i.e. the degree of artificial depth effect), is poorly characterized.

2.2 Creating a full video module, a step-by-step overview

This section is not a summary of the finer details included in the rest of the document. However, familiarizing oneself with general steps will make the purpose of the other sections easier to understand. These steps outline the process from camera setup, to importing to your personal computer, to making adjustments to the stereoscopic properties, to adding video effects, and exporting to various devices.

This document focuses primarily on producing and delivering SBS media (as opposed to the equally common, red-cyan color anaglyph format). However, both methods of 3D display can be inter-converted between each other, even at the time of viewing—if uploaded and viewed on YouTube with appropriate

¹Biddle M, Hamid S, Ali N. An evaluation of stereoacuity (3D vision) in practising surgeons across a range of surgical specialities. *The Surgeon*. 2014 Feb 1;12(1):7-10.

²Schoolman A, inventor. Stereoscopic endoscope arrangement. United States patent US 4,651,201. 1987 Mar 17. The perceived depth from a stereocamera setup is not dependent only on the intercamera-distance, but also the distance to the video subject. The relationship between this distance and the ICD is what produces differences in parallax between the eye feeds—which is where the depth effect comes from.

³Garg AX, Norman G, Sperotable L. How medical students learn spatial anatomy. *The Lancet*. 2001 Feb 3;357(9253):363-4.

metadata (see section 7.2.1). When choosing which method to use, there are some important advantages and disadvantages to consider.

Step 1 Take two smartphone cameras (very ideally Android devices) and mount them with displays facing the subject specimen. Place them in a stable container (a makeshift cardboard container and duct tape will suffice) with front-facing cameras aligned vertically, and both phones placed landscape-fashion, parallel with the ground. Minor deviations from this alignment can be fixed in post-production, but getting it as accurate as possible now will prevent unnecessary cropping at the edges of the video. The cameras themselves can be placed *in parallel*, i.e. the lines of sight do not have to converge at any central focal point (see section 3.2 for why). In this setup, make sure that both phones are connected, wired, to a single laptop.

Step 2 Launch *Vysor*, screen-viewing software for Android devices (<https://www.vysor.io/download/>)⁴. Open both devices to view and interact with the smartphone camera screens through the Vysor application. Some Android devices may not register for file or video data transfer—for best results, use a USB-A to USB-C cable⁵. Use your cursor to open the camera app and press record on both devices via Vysor’s screensharing to your laptop (the time in-between recording does not matter).

Step 3 To later sync the timelines on the footage, create a *slate* with your fingers. Very quickly tap your index together and thumb together in view of both cameras. From here, act out your scene as normal—use probes, props, audio, etc. within view of the object to teach the anatomy of the specimen. When finished filming, use Vysor to end the recording on both devices⁶.

Step 4 Transfer video footage from each phone to your desktop. Rename the footage with any nomenclature you like, but ensure that left-camera and right-camera footage is labeled accordingly. Enable *USB debugging*, and ensure *USB file transfer* is turned on by default in Android phone settings.

Step 5 Download StereoMovie Maker (2014, version 1.30a, Masuji Suto: <http://stereo.jpn.org/eng/stvmkr/>). StereoMovie Maker is a single .exe that accepts video footage, and allows a variety of edits to be made before exporting. By virtue of its age, SMM only accepts .avi or .mpeg files (and usually not the former, either), some of which are nearly fully deprecated or unsupported online. However, the magnitude of functions rendered possible with SMM makes this software invaluable. Frame-by-frame temporal syncing, convergence adjustments, rotating and automatically cropping to adjust the stereo window, color correction, SBS/anaglyph exports, and more are all possible. The next steps will outline how to prepare a file for SMM, and then how to process it for assembly in conventional video-editing software to create a full video module.

Step 6 Download FFMPEG, a fully open-source command-line controlled single .exe file (<https://ffmpeg.org/>). FFMPEG allows you to determine input file paths and output file paths, open, and run the executable from the command line. For readers unfamiliar with the Windows command line, follow through with these steps. Take all of your videos (likely in .mp4 format) and place them in one folder in Windows Explorer. Move ffmpeg.exe into the same folder. From that folder, in Windows Explorer, click on the location or filepath bar (It should read something similar to: This PC / Data (D:) / ResearchProjects...), and type in “cmd”. This will open the command line.

To convert all footage in that folder to an .mpeg (1, not 2, 3, or 4) file format, type the following into the terminal (in a single line), and press Enter. This code snippet will take every “.mp4” file in the directory (dir, or folder), run FFMPEG on it, to convert it to an .mpeg file. Other commands such as -qscale 0 ensure that the quality of the stream is as close to the original as possible. -r 24 sets the framerate of the video at a constant 24 frames per second. This is adjustable, but check what the framerate of the source footage is first. It IS possible to fake a higher FPS on a 20 fps video, but this will not be helpful when managing video quality.

⁴An open-source but less intuitive screen-sharing application to display and control your Android device from your laptop is *scrcpy* (<https://github.com/Genymobile/scrcpy>). *Scrcpy* uses command line, while Vysor comes with a GUI.

⁵Some Pixel 2 devices will not be able to transfer files with a USB-C to USB-C connector, but *will* work with a USB-A to USB-C cable. However, its predecessor, the Nexus 6, can use USB-C to USB-C cables. Try different cables, and enable USB debugging. Follow any other instructions provided with the Vysor software. Purchasing premium is not necessary.

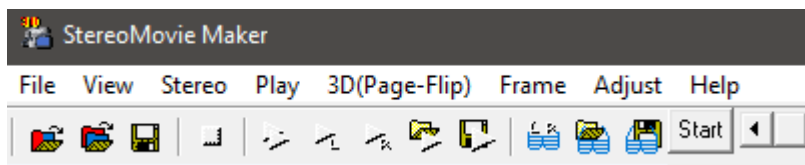
⁶With extended days of filming, keeping two smartphones with open camera apps in a cardboard enclosure next to each other can cause overheating. Do not panic if the phones shut down randomly. Point a desk fan at the setup to dissipate heat, or turn the phones off in-between shots.

```
FOR /F "tokens=*" %G IN ('dir /b *.mp4')
DO ffmpeg -i "%G" -qscale 0 -r 24 "%~nG.mpeg"
```

If only converting one video, the following snippet will suffice. `-i` refers to input; the file extensions/endings are, of course, changeable. When managing voiceovers and recordings, FFMPEG can be used for audio file conversions as well.

```
ffmpeg -i "[insert file name here].mp4" -qscale 0 -r 24
"[desired output file name here].mpeg"
```

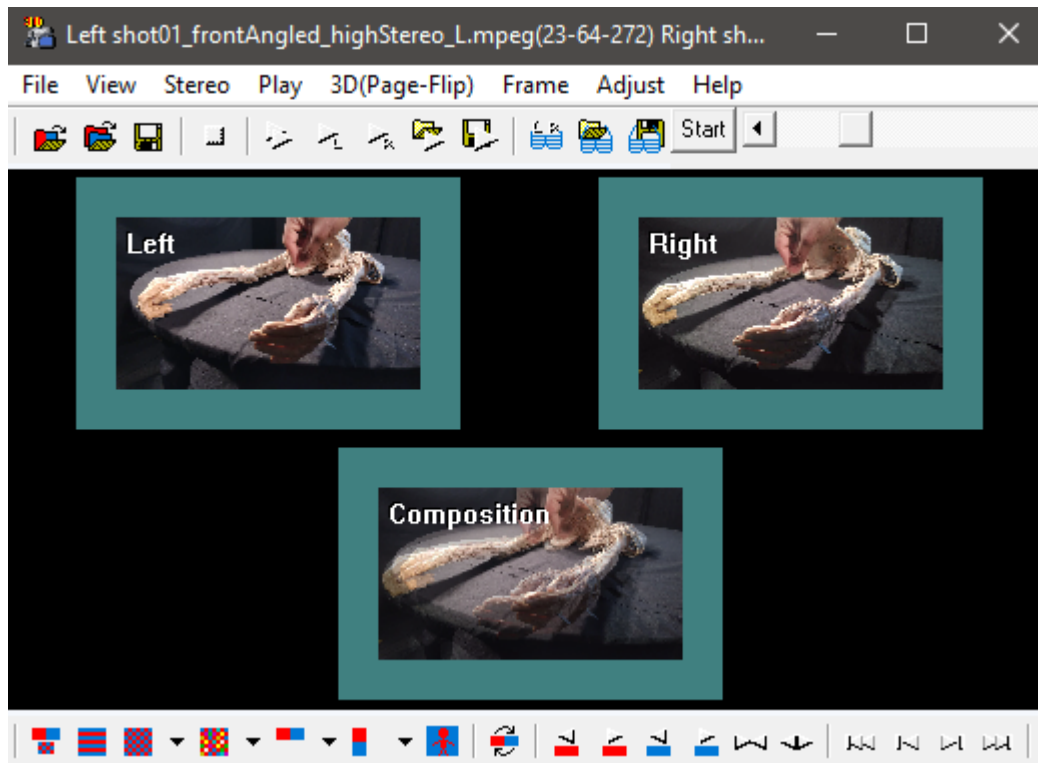
Step 7 With .mpeg files in order, open SMM. Sparse tutorials for StereoMovie Maker are available online, so some discussion on how to use the software is included here. For each set of L/R stereoclips, a few adjustments will need to be made. The clips will need to be temporally aligned, since recording started at different moments. The clips will also need to be rotationally and translationally aligned and/or cropped to ensure the images are converged properly in stereoscopic viewing.



Navigate to File → Open Left/Right Movies, and select your left video, and then the corresponding right video. To temporally align video (do this part first), play both videos with the triangular Play icon on the top icon menu. Stop the videos when one hand enters the frame in one video. Use the red and blue triangular icons on the bottom icon menu to scrub forwards frame-by-frame on the left and right videos independently, to the frame where the fingers slate together in both videos. Navigate to and press Frame → Set displayed-frame to start frame. Your video is time-aligned.

For rotational alignment, several options are available to you. I suggest experimentation with the different settings under the Adjust menu. Meaningful results for SBS video can be achieved by manipulating only the Auto-Adjustment settings. Select “no adjustment of stereo window” (this should be default), uncheck “correct barrel distortion”, and uncheck “H. and V. perspective rotations”. Click OK, and then navigate to Adjust → Auto-Alignment (or, press Alt-K). The image display in the main window should adjust significantly, and the overlap in the “composition” should be sensibly aligned between left and right views.

Color equalization options are sparse but can be tried out View → Auto-Color Adjustment. To export, open File → Save Stereo Movie. Experiment with different options. I resize and resample if necessary to 1920x1080. Export in side-by-side format. Do not bother exporting as full frames—pick a reasonable compression format to save filespace, and make video management easier. When ultimately displayed on a phone, the end resolution will be half of a conventional smartphone screen. Repeat these steps (try to stay consistent) with all of your other clips from the day’s filming.



Step 8 Any video editor of choice can be used to assemble temporally and perspective-adjusted SBS stereo footage exports from Step 7. For the Brachial Plexus Neuroanatomy video produced at McMaster (www.github.com/malyalar), Adobe Premiere was used for robust VFX integration, with or without Adobe After Effects. Add text effects, label lines, voiceovers, transitions, and more appropriately. Details on how to create text or line effects, while 1) not intruding on the stereo window and 2) maintaining a sense of depth in the VFX itself, are covered in section 4.1 and 4.2.

Step 9 Export finished SBS video in any file format supported by the chosen video editor. A common container format playable on a variety of smartphones is an .mp4 using the H.264 codec. To upload to YouTube and allow the site to 1) recognize the video as a stereoscopic SBS 3D video, and 2) allow it to convert to anaglyph, a few extra steps will need to be undertaken. See section 7.2 for how to inject the code for the YouTube-compatible metadata, and section 7.3 for what edits need to be made such that the conversion maintains a comfortable, convergeable stereoscopic 3D experience. No extra software is necessary; FFmpeg will suffice for metadata injection.

Otherwise, any video player on any smartphone can be used to open the video, and a well-produced video could reasonably be played on any type of Cardboard headset. MXPlayer (freemium, ad-supported, Android application) will allow scaling of the video to take up less or more screen real-estate, acting somewhat as a way to adjust the inter-pupillary distance between videos. Some bulkier Cardboard-compatible headsets such as the *ETVR* allow for IPD between the lenses and the forward travel of the lenses to be adjusted as well. Such headsets may produce better results for some viewers, but generally there is more to be fixed in the video creation process than the video display process. Diplopia (double vision) is common, and difficult to surmount.

3 Constructing a stereoscopic photoshoot environment

While setting up a photoshoot can seem trivial, there are several details that contribute significantly to the quality of the final product. Not every error made at this stage is fixable at post-production, and with stereoscopy in particular, there are extra factors to consider in your arrangement of the apparatus and scene. Nevertheless, prototyping is easy, and the tolerances of these various factors are generally

somewhat forgiving—until professionalism and versatility among wide audiences with differing viewing needs becomes critical.

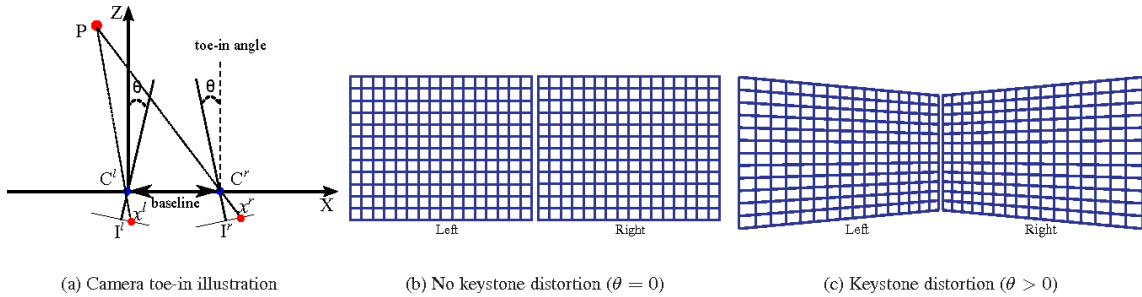
3.1 Dual smartphone camera apparatus

In the interest of democratizing not just the distribution and consumption of the stereoscopic media, but also its production, this document seeks to use two consumer smartphone front-facing cameras of average quality (in 2018) to create the video. Both devices do *not* need to be identical. While there are intrinsic differences between device cameras with respect to focal length, aspect ratio of sensor, and more, most android devices and most camera software will use the sensors located within in similar ways that produce comparable results with any device. Other intrinsic devices between phones when using two separate models with subtle height differences, affecting the transational positioning of the stereo window. You may choose to account for this difference with some adjustments to the apparatus used to hold the phones onto each other.

3.1.1 Toe-in vs. parallel camera setup

Stereoscopy works by taking two basic captures of a scene, and then presenting the views separately to a viewer’s eyes. If the two basic captures of the scene (the camera feed) were to mimic a person’s eyes completely, then the cameras may be tilted inwards towards the object. However, when doing this, the camera creates an effect where as each eye observes an object into extreme fields of view, the object also becomes closer. For the right eye (red), the left side of the visual field is artificially smaller. For the left eye (cyan) the right side of the visual field, in the digital wall, is small. This visual error is referred to as keystoneing, and while it is never an issue in real life, especially when dealing with close-subject stereoscopic media, it can significantly impact the ability of the viewer to converge the image properly.

Hence, the best method is to use a parallel camera setup, where both cameras point infinitely forward with no plane of convergence. This completely avoids keystoneing⁷, and while there are technically some realism issues regarding the “proper” perceived depth of stereo (e.g. when the eyes accommodate for close objects, they attempt to accommodate realistically, but all objects are actually at the plane of the screen, causing discrepancy), the ease of successful convergence with a wider array of objects and distances makes parallel camera setups the go-to choice. Using smartphones in particular, it is also easier to mount them in parallel than tilt them carefully to achieve a shared focal point precisely X inches away from the phones.



3.1.2 Focal length to distance and inter-camera distance

The actual quantity of ‘depth effect’ produced in the stereoscopic shot is a ratio of the inter-camera distance over the distance to the object(s) with depth in question. With wider differences in inter-camera distance, more lateral views of the object are obtained, increasing *static parallax* perception, and providing the eyes with greater depth information (to a degree, until convergence can no longer be achieved). Different individuals respond differently to changing inter-camera distances⁸.

⁷Liu F, Niu Y, Jin H. Keystone correction for stereoscopic cinematography. In 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops 2012 Jun 16 (pp. 1-7). IEEE.

⁸Tested independently with a computer generated stereoscopic rendered video, where the cameras slowly move closer together while investigating the scene. Some viewers do not notice the changing ICD, while other views experience obvious

In practice, this indicates that there may be not a specific ICD, but a specific ratio for optimal comfortable levels of depth in a scene. This ratio ranges from 1:20 to 1:30 for ICD:Distance to Object. Significant excursions from this ratio are reasonably well tolerated in a variety of audiences. However, when wanting to test extreme circumstances, very high ratios (hyperstereo) can produce diplopia.

3.1.3 Hyper and hypo-stereopsis conditions

The plausible benefit of hyperstereo is increased depth perception due to increased information from increased static parallax. In practice, hyperstereo images tend to converge poorly, and provide marginal meaningful depth information, while being harder to manage videographically, especially come time to apply video effects and text effects to the screen. Hypostereo or normal stereo images are more comfortable and still produce convincing 3D effects—convincing enough that subjectively it is very difficult to evaluate and differentiate between hyperstereo and hypostereo scenes consistently for many subjects. That being said, learning difficulties do not necessarily equate to worse learning, and so it is a deserving area of study whether hyperstereopsis produces a meaningful learning effect regarding spatial anatomy.

When setting up the camera mount, if looking to modify the stereo level, ensure that the phones can slide laterally easily to move the cameras closer or farther apart.

3.2 Adjusting for inter-pupillary distance

Differences in interpupillary distance, at least on population scales, DOES produce differences in depth perception. Those with smaller IPDs who still have normal binocular vision tend to resolve depth a little more poorly than those with larger IPDs. More importantly, differing IPDs complicate the process of projecting 3D information onto a 2D window⁹. Despite being able to recognize these problems, however, little can be done to effectively resolve them. Certain Cardboard compatible headsets such as the ETVR allow for lateral adjustment of the eye lenses, which when combined with corresponding lateral displacement of the video on the smartphone display, may allow the same video to be shown with equal opportunity for vergence to both people with close eyes and far-apart eyes.

However, such methods for fine adjustment for IPD differences are almost universally inconvenient, and likely not worth sinking time into. Using an inoffensive level of stereopsis that is comfortably converged by a variety of IPDs in the first place seems to be a more effective approach when considering how best to disseminate VR anatomy learning experiences. Inconvenience is a massive factor for poor uptake of VR technology, inside or outside educational realms.

3.3 Resolution and framerate

Temporally syncing footage is most easily performed when both cameras are recording at the same framerate. Aside from a subjective minimum of 20 frames per second for smooth footage, using a third-party camera app such as OpenCamera, an open-source Android camera application written by Mark Harman (<https://opencamera.sourceforge.io/>), the recording speed for either the front-facing or rear-facing camera can be set pre-footage under Settings (Large gear icon) → Video settings → Video framerate. The set frame rate may be approximate, and is not guaranteed to be achieved. The video recording may fail if the frame rate is not supported.

At this same screen on OpenCamera, video resolution can also be set. To avoid resizing issues when placing footage side-by-side in your video editor, try to use the same aspect ratio and resolution in both cameras. However, if frame rate and resolution are proving difficult to modify appropriately, both can be modified in post-production (see sections 6.2 and 6.3).

There will usually be subtle but perceptible differences even when trying to match frames. Artistically, in SBS stereoscopic footage, approximate temporal matching is acceptable for close footage. The

changes in the stereo depth (but are not necessarily nauseated by it).

⁹I highly suggest reading through: Wann JP, Rushton S, Mon-Williams M. Natural problems for stereoscopic depth perception in virtual environments. *Vision research*. 1995 Oct 1;35(19):2731-6.

tolerances are within very small fractions of a second, and the less the better, but when slating footage (see section 2.2, Step 7), you may notice that the fingers are not precisely in the same locations in any frame. This is not intractable, and the stereoscopic effect will still be convincing, assuming everything else is within order.

3.4 Lighting, contrast, shadows

In the absence of stereoscopic information, lighting, contrast, and shadow angle or length are strong ways to convey depth in a scene, and contribute significantly to the impression of depth¹⁰. These factors continue to inform depth perception even after re-including stereopsis in the viewing experience, as indicated by placebo stereopsis¹¹.

Color, lighting, and shadows can be manipulated by adjusting the smartphone camera, the scene lighting itself, or can be done in post-production. There is little to state other than to ensure that shadow and lighting information corroborates information obtained from stereopsis. This makes integration and convergence easier for viewers. A standard three-point lighting setup, without washing out any structures, is more than appropriate. More importantly, stay consistent between shots: keep lighting similar between angles, and avoid applying different color grading filters at the time of filming with the camera application.

4 Labeling and artificial/digital stereoscopic 3D VFX

4.1 Artificial depth for text

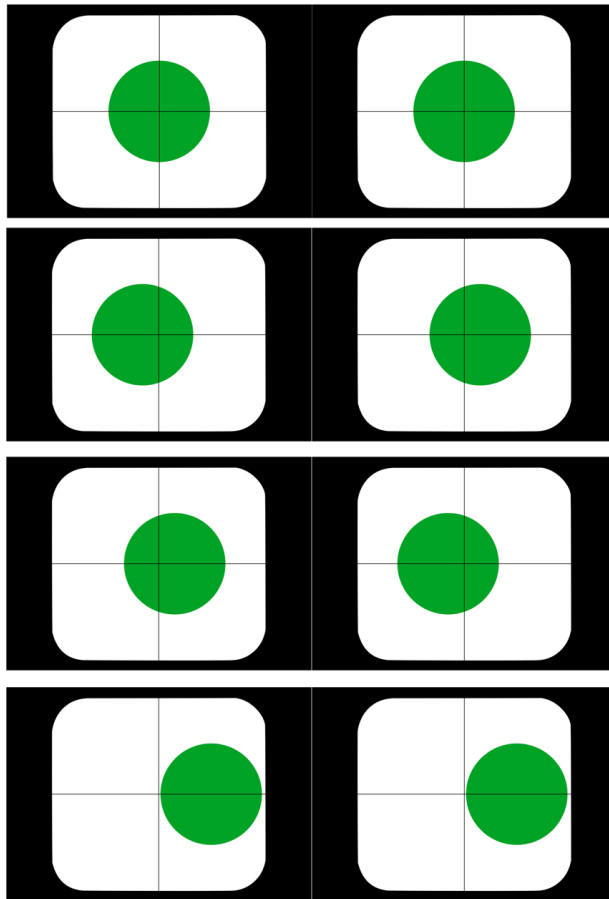
The following figure is a powerful reference for how differentially placed VFX at different parts of the stereo window can produce different effects. For text in particular, proper convergence is crucial. I suggest using a low-stereo scene with minimal requirement for eye vergence, and placing text as close to the the centre of the screen as possible. I also suggest matching up corners of the lettering with selected “landmarks” on the videoed specimen. E.g. the b in “brachial” should be at the acromion in both left and right views, and the s in “plexus” should be at the deltoid peak in both left and right views. This will NOT produce a depth effect for the text (fourth option in the figure, same plane as the cross hair), but the text will appear to float due to overlapping all other parts of the video.

Using high contrast, VR friendly fonts is key. Sans serif typesets with large font sizes and a thin black stroke outline can help avoid poor contrast from text to specimen. Overlap is a crucial part of the perceived depth effect and diplopia is common enough: the more clear the depth information, the less likely that double vision will occur.

Animating the text effects’ entrance and exit may improve the appeal of the final product.

¹⁰Hubona GS, Wheeler PN, Shirah GW, Brandt M. The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. *ACM Transactions on Computer-Human Interaction*. 1999 Sep 1;6(3):214-42.

¹¹Placebo stereopsis is colloquially known as the *Van Hare Effect*, where a single monoscopic image is displayed on both the left AND right sides of an SBS display. Subjects often report perceiving stereoscopy where none can exist. See https://en.wikipedia.org/wiki/Van_Hare_Effect



These are identical pairs. The circle in the left frame and the right frame are in the same location. When viewed in 3D, the circle will be on the same plane as the black mask and the cross hair, and will be centered in the frame.

In this example, the left circle is shifted to the left and the right circle to the right. When viewed in 3D, the circle will be behind the black mask and the cross hair. And will be centered in the frame.

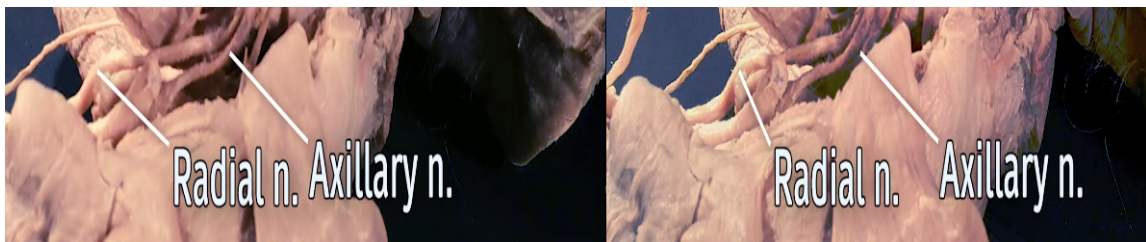
In this example, the left circle is shifted to the right and the right circle to the left. When viewed in 3D, the circle will be in front of the black mask and the cross hair. And will be centered in the frame.

In this example, both circles are shifted to the right. When viewed in 3D, the circle will be on the same plane as the black mask and the cross hair. The circle will be on the right side of the frame.

4.2 Lines, moving towards or away from the screen

Text labels are simple objects which stay in one plane relative to the stereo viewing plane. However, with callout label lines, another question arises: should the labels appear to be moving into the scene, onto the specimen from the text? In the cropped screenshot below, a label of the radial and axillary nerves are visible, both from the left and right eye feed. Both text labels are in one stereo plane. The label line for the axillary nerve is the same for both left and right images. However, the label line for the radial nerve is angled differently in the left eye versus the right. In both cases, the line travels from the left 90-degree peak of the “R” in *Radial n.*, and points towards the same spot on the radial nerve itself. However, because of static parallax, the line is angled differently. When viewed in Cardboard, an impression is given, with minimal double vision, that the line is traveling directly into the specimen, into the screen.

This effect was produced very approximately and haphazardly, only by matching the ends of the line with two separate landmarks—a chosen point on the text, and on the specimen. Other such dynamic VFX can be similarly produced by taking advantage of the stereo reference inherent to the specimen itself. Experimentation with the maximum tolerances for such FX can be helpful.



5 Shot selection and composition

5.1 Depth differences in a single frame

The image below is an anaglyph of two L/R wide angle shots of a brachial plexus specimen, seen from the inferior aspect. I will refer to the distance between the red and cyan anaglyphs for any part of the image “divergence”. Divergences within this single specimen are very large compared to close-ups in other parts of the brachial plexus module. Observe that the divergence at the hands dwarfs the divergence at the ribcage and face. Divergence (or convergence) can be modified in post by moving the two L/R stereo images horizontally towards each other. However, if the images here were to be moved closer, while the image at the hands would converge, the image at the ribcage and head would *diverge* further. This is also especially visible with the text. Text pointing out the biceps muscle is comfortably converged, but text pointing out the skin of hand is uncomfortably diverged.

In specimens or angles with a great deal of depth difference in a single shot, diplopia or double vision caused by poor convergence is common. The eyes cannot converge all parts of the image due to these large disparities in the VR stereoscopic environment. When selecting shots and scene compositions, consider selecting angles where all parts of the image are roughly in line with the intended stereo plane.



5.2 Depth difference between subsequent shots

If an object is perceived to be at a certain distance, the eyes can focus rotate orbitally, and focus their lenses to that distance (processes called convergence and accommodation respectively). In stereoscopic imaging however, the light reaching the viewer’s eyes from close-by virtual objects will still come from the actual screen, which is rarely at the same perceived distance as the object in the scene. And so, the eyes will focus on the wrong plane, and the image may appear blurry.

These processes compound over a several-minute long video. From shot to shot, angle to angle, different scenes rendered with different IPDs, or the same IPDs but with different distances to the subject, will have their respective static parallaxes, which can each vary significantly. Subjectively, going from one shot to the other can thus be a painful experience, as your eyes quickly re-converge and re-accommodate at new points. Hence, care should be taken when considering the shot order of the exact IPD/distance ratio of each shot. The change from shot to shot should be minimal. A transition from a far shot to a close shot with no appropriate change in IPD can produce difficulties for viewers.

5.3 Moving and panning shots

Motion parallax is another important aspect of depth perception. With *Google Cardboard*, the user is able to rotationally explore their environment using combined information from the accelerometers and gyroscopes in most consumer smartphones. However, head-tracking with *Cardboard* is currently impossible, and with the SBS stereogram characterized in this document, there is zero ability to *dynamically explore* the anatomical subject¹².

Hence, it may be more crucial to include moving or panning shots in the final video learning object. Smooth, professional pans are hard to achieve without professional camera stabilizing equipment, and post-production for wobbly footage is difficult, often severely distorting the image. The close distance from cameras to subject make complex moving shots extremely difficult. However, with careful tripod work, smooth and slow panning shots are doable. Use a rubber band on the tripod stand to pull with a consistent speed. If moving shots are not possible, use a variety of stills to provide the viewer with multiple angles on the same specimen¹³.

Using a series of stills may be off-putting for a learner expecting moving video footage. To create movement, either move the object manually while the camera setup remains still, use a probe to point at points of interest on the specimen in front of the cameras, or create a *cinemagraph*. A cinemagraph is short looped footage where all major movement is removed, keeping only easily-looped, small-scale movement. If your specimen is plastinated or wet, small portions and threads and structures may be easily mobile. Point a strong desk fan at the object to cause mobile parts to vibrate or move back and forth subtly. These small movements are noticeable and appealing in footage to create a sense of liveliness, but are also small-scale enough to be easily played on repeat without drawing attention.

6 Color grading and post-processing

6.1 Equalizing color grade

Automatic, albeit low-quality color equalization in StereoMovie Maker with Ctrl+K is outlined in section 2.2, step 7. Automatic color matching methods are also available in the most recent versions of Adobe Premiere and the Premium Studio version of DaVinci Resolve (general purpose video editing software with an emphasis on color correction utilities). Otherwise, color matching will have to be performed manually with creative use of masks and color curves. While the best color matching is no color matching (because hopefully your smartphone cameras will have been of the same model), color grading is almost always warranted to achieve a more professional cinematic look.

With anatomy specimens in particular, color interpretation may be crucial for differentiating structures. The role of contrast and lighting has already been discussed in this document, but aesthetically, increasing the warmth of skin tones in specimens, and cooling background objects can vastly improve the appeal and watch-ability of a shot. Color correction is a complicated art; take a look at some online tutorials on how to create a cinematic look in your footage.

That said, there is a lot of forgiveness for lack of color coordination between the left/right images. Changing one video feed to full greyscale, for example, will not place undue stress on the viewer. The image will appear to multiply color between the greyscale and color images, and converge appropriately.

6.2 Lateral borders

With reference to the image below, when producing SBS stereoscopic media, observe that the field of view from the right camera will necessarily capture parts of the environment that are NOT covered by

¹²Luursema JM, Verwey WB, Kommers PA, Annema JH. The role of stereopsis in virtual anatomical learning. *Interacting with Computers*. 2008 May 1;20(4-5):455-60.

Dynamic exploration of a subject is particularly important for learners with poor *visuo-spatial ability*, i.e. “the ability to form, retrieve and manipulate mental representations of a visuo-spatial nature.”

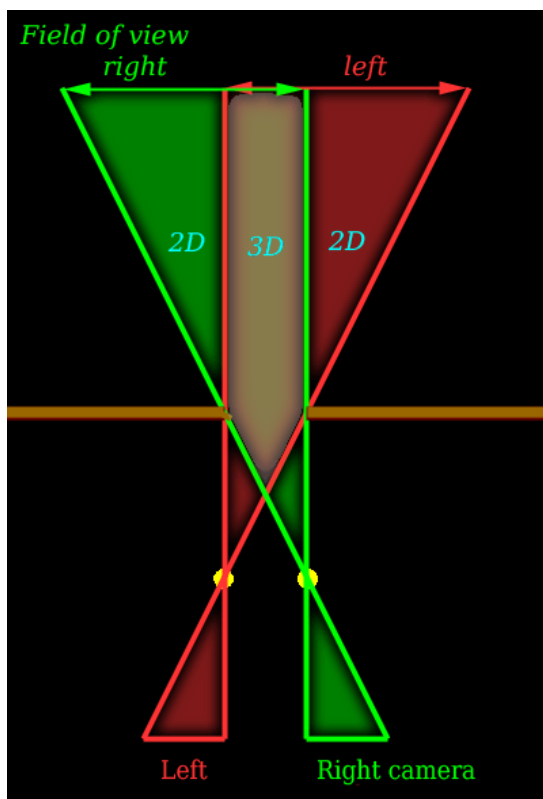
¹³Viewing specimens from multiple angles improves spatial understanding.

Garg AX, Norman G, Sperotable L. How medical students learn spatial anatomy. *The Lancet*. 2001 Feb 3;357(9253):363-4.

the left camera, and vice versa. Hence, when viewing the leftmost and rightmost edges of your field of view in an SBS-VR headset, those images are non-stereoscopic: only informed by one eye. Perceptive viewers may also notice double vision in those regions, interchanging between the eye’s view of the object, and darkness. This is often distracting, since the most important parts of the specimen tend to be in centre frame.

To resolve this issue, consider placing black borders on the non-stereoscopic edges of the SBS field-of-view. This will reduce the viewable area in the video, but can also focus attention onto the center of the specimen, where stereoscopic fidelity is greatest. In some headsets, especially headsets made of actual cardboard, there also tends to be a significant visible chromatic aberration at the edges of the field of view, both horizontally and vertically.

This aberration stems from image distortion by the lens as it focuses a rectangular image into a circular pinpoint (your eye). Individual fisheye distortion of the SBS video feeds may help marginally, but chromatic aberration is common in any cardboard VR experience—just more pronounced with smartphone produced VR footage. Aside from using bars to cover the edges of the FOV, also consider applying a liberal vignette to darken out the edges of the screen. This may make structures appear more hazy, but may also again focus attention to the centre, where visual fidelity is greatest.



7 Publishing to the web

7.1 Differences between screen hardware

The screen differences between the Nexus 6P (2015) and the Pixel 2 (2018) will be used to illustrate certain problems with smartphone displays. Newer is not always better; resolution is paramount, and older devices can have better resolution, especially with larger screens. The Nexus 6P uses a Samsung AMOLED display with a “Wide Quad HD” (2560x1440 pixels) resolution at 518 ppi. This wide resolution is particularly useful if you take footage at 1920x1080 and proceed to render SBS footage at 3840x2160 (consider resizing to the average or target device resolution). The Pixel 2 also uses a Samsung AMOLED display, but with a standard 1920 pixel resolution at 538 ppi. While pixel density is higher, fine details

are often washed out entirely on the pixel display. However, large structures most crucial to the learning of anatomy are maintained on either screen.

Size differences between screens may also complicate IPD-related issues. Using MXPlayer (freemium Android application) will allow you to resize the video and adjust image size (and thus IPD distance) on demand with simple gestures.

7.2 Publishing to YouTube (with metadata)

For H.264-encoded video in an MPEG-4 container, open cmd in the correct Windows Explorer directory (with FFMPEG in the directory as well), and type in:

```
ffmpeg -i input_file.mp4 -vcodec libx264 -x264opts frame-packing=3 output_file.mp4
```

Looking online for frame packing solutions, you may find this code snippet. This will present an invalid argument—ignore it.

```
ffmpeg -i input_file.mkv -vcodec libx264 -x264opts  
"frame-packing=3:frame-packing-interpret=1:frame-packing-quincunx=0:  
frame-packing-grid=0,0,0,0" output_file.mp4
```

7.3 Conversion between SBS and color anaglyph

After injecting YouTube-compatible SBS metadata into your .mp4 video using FFMPEG, a regular upload to YouTube will suffice. Depending on the length, resolution, and size of the video, YouTube will take more or less time to process it. It will all be done automatically, however. When viewing the video on a smart phone's YouTube app, click the gear icon and note that you can switch between Cardboard compatible SBS view, and regular anaglyph view.

However, conversion is more than just coloring one image red, the other blue, and then superimposing the two images on top of each other. When automatically superimposing the images, the centres of both the left and right videos matter. The relative positions of the video feeds on the real estate dictate the level of convergence. Convergence(s) appropriate for SBS film will not necessarily work for anaglyph videos, and vice versa. Hence, each individual shot needs to be re-tailored in StereoMovie Maker to acquire the best fit for anaglyph viewing. Horizontal convergence adjustment, and automatic cropping of excess aspect, can be done in the Adjust → Easy Adjustment menu in SMM.