Movidius

Warning: Objects on the screen are closer than they appear!

White Paper

October 2010

Myriad 3D

1 Overview

This white paper provides a background on mobile 3D technology and how Movidius' Myriad 3D platform is enabling handset device designers to incorporate high quality mobile 3D video capture and playback solutions for global markets.

1.1 Eruption

If the impending surge of mobile video traffic resembles an approaching tsunami, then 3D technology on a mobile device is a deep rumbling volcano ready to erupt. A large ash-cloud of 3D glasses, displays and depth cues is preparing to blow its way around the mobile world.

This paper looks at the triggers for this explosion – the push of Hollywood 3D content to attract more audiences into cinemas coupled with the pull by TV OEMs to upgrade from HD to 3D TV sets together with the move towards mobile 3D enabled by the Movidius 3D video processor.

Warning: Objects on the Screen are Closer than they Appear!

1.1.1 The Challenges of Mobile 3D

Since the Consumer Electronics Show (CES) in early January 2010 there has been an explosion of interest in 3D TVs. The first live 3D sports broadcast followed soon afterwards on 31st January by Sky TV.

James Cameron's Avatar blew away audiences around the world with the 3D effects. Suddenly the 3D movie viewing experience became an enjoyable and absorbing experience rather than a headache-inducing collection of not-quite-right images. Similarly in the gaming environment, the Nintendo 3DS launching in 2011 uses a glasses free 3D touchscreen.

Just as the trend in the mobile industry is stampeding towards greater video production and video sharing, can we assume that it is only a matter of time before 3D mobile phones become mainstream?

Perhaps, but certainly the mobile 3D ecosystem needs to tighten up and address some significant technology challenges. Unquestionably, the technology advances in the 3D display industry have enabled a range of 3D solutions which was not previously possible. From the user experience perspective the glasses-free (auto-stereoscopic) use-case for the 3D mobile device has now become a reality.

The availability of 3D content will be a major motivator for the 3D user. Already major TV sports channels such as Sky 3D and ESPN 3D are broadcasting live matches in 3D. Many of the viewers watch in public venues or bars but the push is on for the home 3D system. The 3D movie production is gaining some momentum too.

3D content generation is still very small compared to 2D content, however, and a clear missing piece of the puzzle is user-generated-content created by the general public. Enabling the average user to produce 3D content locally on mobile devices will increase exponentially the 3D content appearing in the mainstream.

Finally, playing standard 2D video streams and previously captured user-generated videos on the user's newly acquired portable device's 3D screen can be greatly enhanced by converting the video to a 3D video in real-time. Quality user experiences will depend on the quality of the conversion algorithms used and the realistic 3D scenes that are created.

Finally, the 3D industry has to shake the image that 3D is synonymous with headaches and a tiring viewing experience on the eyes. This will only happen through using the correct technology to deliver a premium quality 3D experience for the user, free from artifacts that show themselves as a painful viewing session.

Intimately understanding the 3D production process is the first step to compensating for these inaccuracies and delivering a robust and enjoyable 3D experience.

1.2 Mobile 3D User Scenarios

There are many mobile scenarios that can employ 3D technology*. The table below shows a portable device with an immersive 3D experience.

3D User Scenario	Technology
3D Still Image Capture	Dual Camera Capture / 3D Image Encoding
3D Movie Capture	Dual Camcorder Capture / 3D Video Encoding
3D Live Preview	Real-Time Live 3D Preview
3D Image Playback	3D Image Decode and Playback / 3D Zoom and Pan / 3D Convergence Adjustment
3D Movie Playback	3D Video Decode and Playback
2D Movie Playback with Real-Time Conversion	Real-Time Automatic Convergence 2D-3D Conversion
3D Gaming	Conversion of 2D Games to 3D
3D Sharing	Upload and 3D sharing / 3D-2D Conversion for 2D Compatibility

**Movidius Myriad 3D solution offers all of these features.*

2 How to Create a 3D Image

Solving the problems with 3D and delivering a credible and quality user experience begins with understanding the technology and what makes a good and bad 3D user experience.

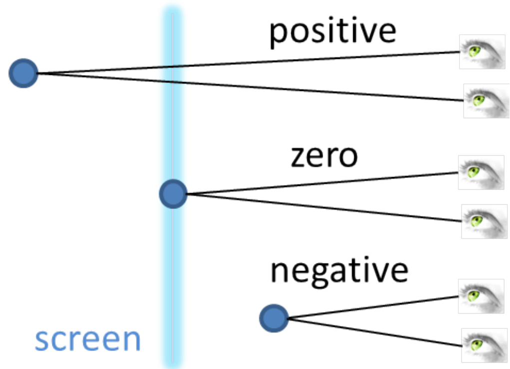
2.1 How is a 3D Image Formed?

A 3D image or video creates a more real life scene by giving the viewer the perception of depth of objects that are displayed on the screen. The underlying mechanism is the introduction of parallax between the left and right images.

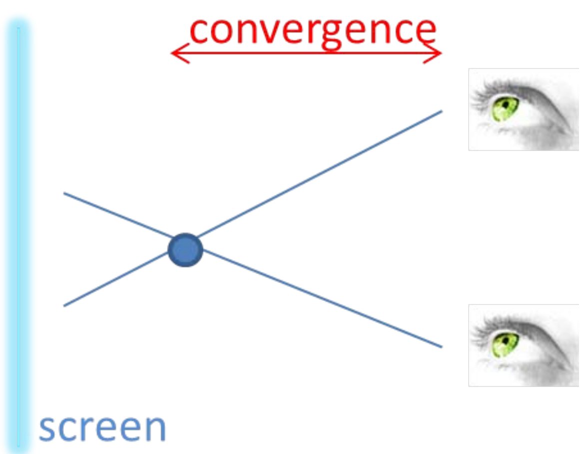
2.2 Parallax

In a 2D video or still image all objects appear on the screen surface – known as the zero parallax plane. Although there is no depth information in the content, the user infers the depth information from a series of cues: perspective (smaller objects must be further away), lighting and shadows, occlusion (near objects block the view of far objects), texture, and relative motion (far objects appear to move slower than near objects).

In a 3D video or still image objects appear to be either in front of the screen, on the screen or behind the screen, giving the perception of depth.



2.3 Accommodation and Convergence



During perception of any real life 3D scene two things of note occur within the human visual system:

1. The eye naturally *accommodates* or changes shape to focus on objects further away. This adjustment ensures that the focal length of the eye matches the distance to the object you are looking at.
2. The eyes also move so that their axis' direction changes, so that they *converge* on a point that is at the distance of the object under focus.

On a 3D display however there is a disagreement between accommodation and convergence. This can often lead to some uncomfortable viewing experiences for users if it is not controlled in an optimum way.

2.4 How a Bad 3D Image is Formed

We are all familiar by now with the 3D demonstration booths at trade shows which range from the stunningly impressive (think Avatar) to the nauseating where fatigue hits your eyes after a 2 minute sample. Understanding the factors which lead to a poor 3D rendering will help us avoid them.

Inferior Quality Convergence Algorithms will produce 3D scenes which will have too great a depth range, making the scene become unrealistic but also but leads to unnecessary strain on the eyes as you try to focus and make sense of the imagery.

Finally some people are actually stereo blind (between 4 and 10%) and will not see the 3D effect. Although for some people, this is a treatable condition. [7]

2.5 3D Safety

Safety is of paramount importance for any consumer product. 3D should be safe, enjoyable, and fun to watch.

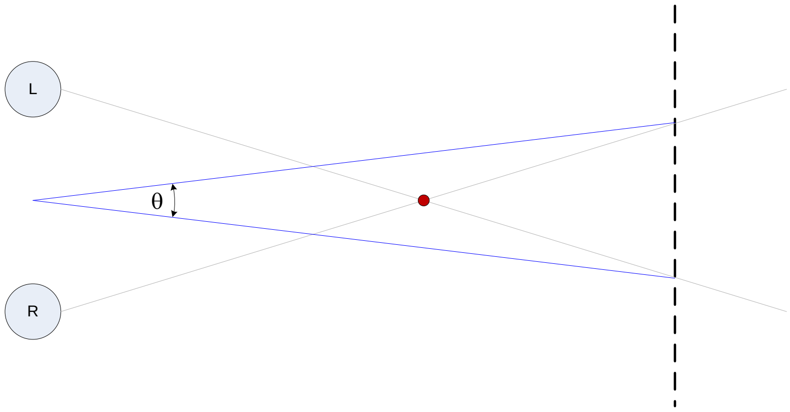
2.5.1 Retinal Rivalry

Retinal rivalry occurs when there is an unnatural disparity between the two images observed with each eye. This can cause viewer discomfort over longer periods of time. Resolving differences such as these is a process known as rectification which is covered in detail in a later section.

2.5.2 Excessive Parallax

As noted earlier, the eye must accommodate and converge correctly so that an image may be correctly viewed as 3D. If one of these two items fail then either double vision or blurred content will occur, both of which are uncomfortable and unpleasant.

Some recommended safety limits have been proposed, and limits have been recommended in literature, such as maintaining a parallax angle (θ) less than 1.6 degrees have been proposed [5]. A good convergence algorithm can determine the depth range of the scene, and maintain parallax within comfortable limits in most cases.



2.5.3 Summary and Key Recommendations

- Choose 3D capture technology with
  - Top quality rectification to limit retinal rivalry.
  - Top quality convergence to limit excessive/bad parallax.

3 Review of 3D Platform Technology

3.1 3D Display Considerations

The mobile 3D use case has some very specific advantages over home 3D viewing when it comes to display technology. The mobile 3D user will be the single user viewing the mobile device from a certain distance and angle. The user is free to move the mobile device to the position that best delivers the optimum 3D viewing experience. This is not always possible in other scenarios where there are multiple viewers.

Also the mobile 3D use case has the advantage of not having to use 3D glasses. This gives the scenario some powerful momentum when migrating users from 2D to 3D technology.

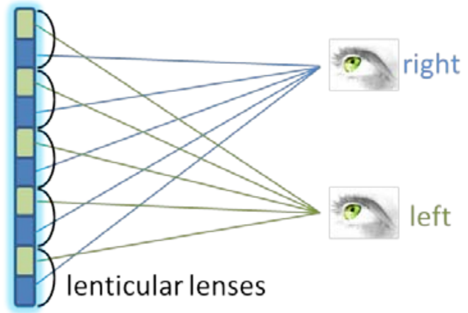
Several types of 3D displays are possible.

3.1.1 Lenticular Lens Arrays

Lenticular lens arrays are possible by overlaying a Lenticular lens array on top of an existing LCD solution. The lens pitch is specific to the pixel pitch of the display.

When well executed this gives quite a good quality 3D experience.

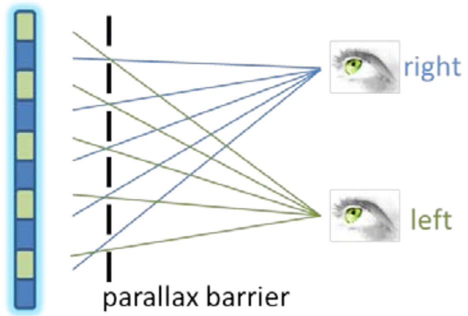
The key limitation is that this is not electronically switchable. It must be either persistently present, or physically added and removed.



3.1.2 Handheld Parallax Barrier Displays

This technology is popular at the moment. A parallax barrier is put in front of LCD at an offset to give good parallax at a particular user distance. The parallax barrier may itself be an LCD which allows for an electronically switchable component. This is ideal for a mobile handheld scenario: full resolution 2D combined with excellent 3D at half horizontal resolution. This is the very much the technology of today, and is used in the Nintendo 3DS.

For this type of display vertically interleaving of content is required just prior to display. This is supported by the Myriad 3D platform.

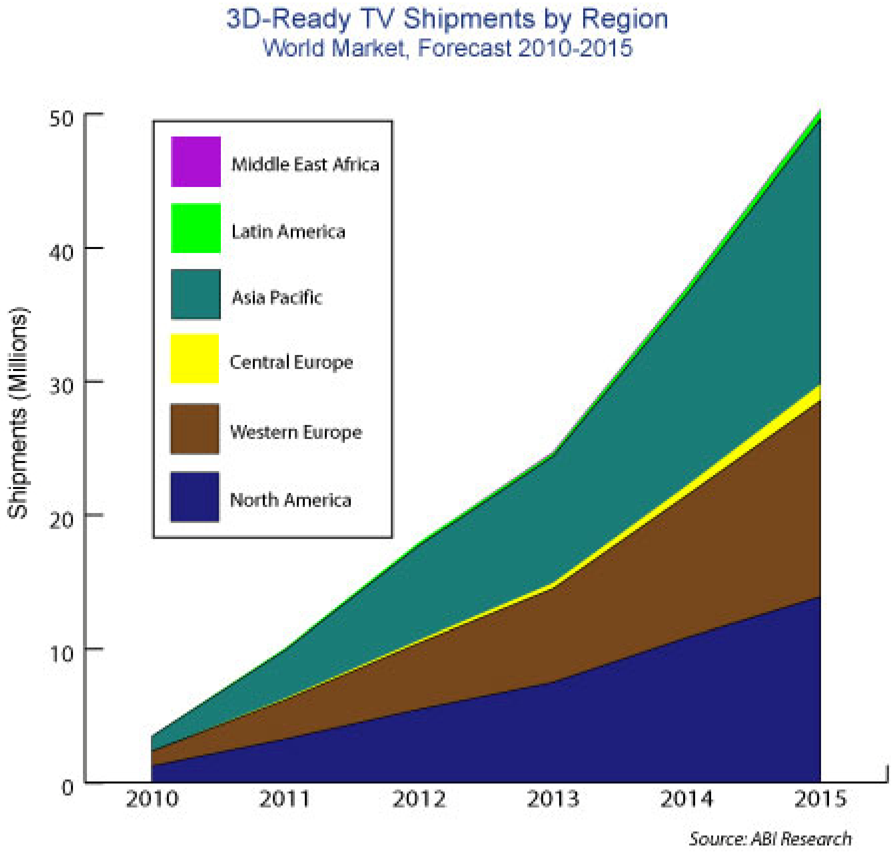


3.1.3 Handheld Backlight Switchable Displays

This type of display utilizes a directional backlight behind the LCD. Frames are shown sequentially at full resolution for right, left eyes in turn with different backlight directions. High frame rates are required for this to work well, and fast loading of frames to the LCD is also required. The technology offers good 3D quality at full resolution. Given stringent interface requirements for high resolutions, this technology is likely to be more popular in the coming years.

3.1.4 HDMI Based Display

With HDMI 1.4a a variety of 3D transmission formats are supported. This allows handheld devices to drive 3D televisions directly. ABI Research's October 2010 report states that market growth in 3D television will accelerate in 2013 and shipments of 3D TV sets will approach 50 million in 2015.



3.1.5 Anaglyph 3D Display

It is also possible to show 3D content on a 2D display using anaglyph technology. This requires the user to wear color glasses, such as the red-cyan glasses used in the '80s. Today some advancement using different color pairs are possible offering less retinal rivalry and less color de-saturation.

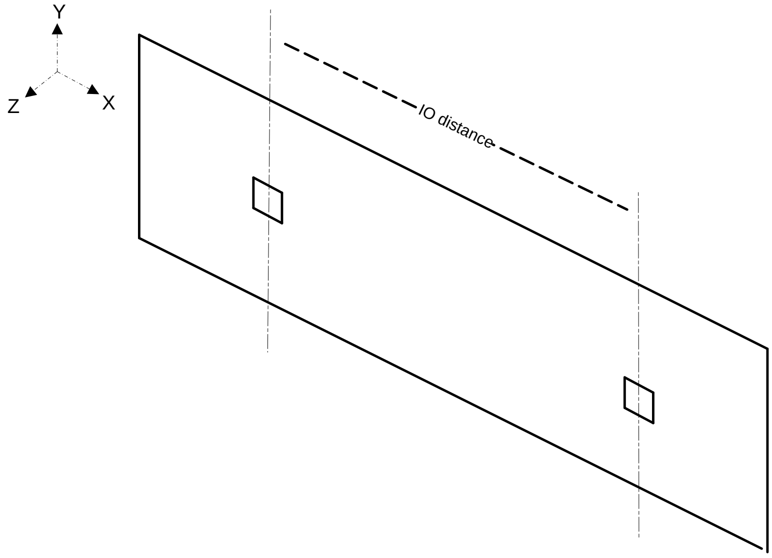
3.1.6 Summary and Key Recommendations

- Handheld devices will be auto-stereoscopic and so not require glasses.
- Being 'glasses free' will be a significant driver for the handheld 3D market.
- Parallax Barrier technology likely to be most popular auto-stereoscopic technology in the near term.

3.2 Camera System Considerations

3.2.1 Overview

In constructing a stereo camera system there are several variables which are possible. As a general rule camera axis should be co-axial, and the Z-Axes should be parallel, and there should be no significant Y offset between the cameras.



To recommend a particular distance between two cameras gives several tradeoffs.

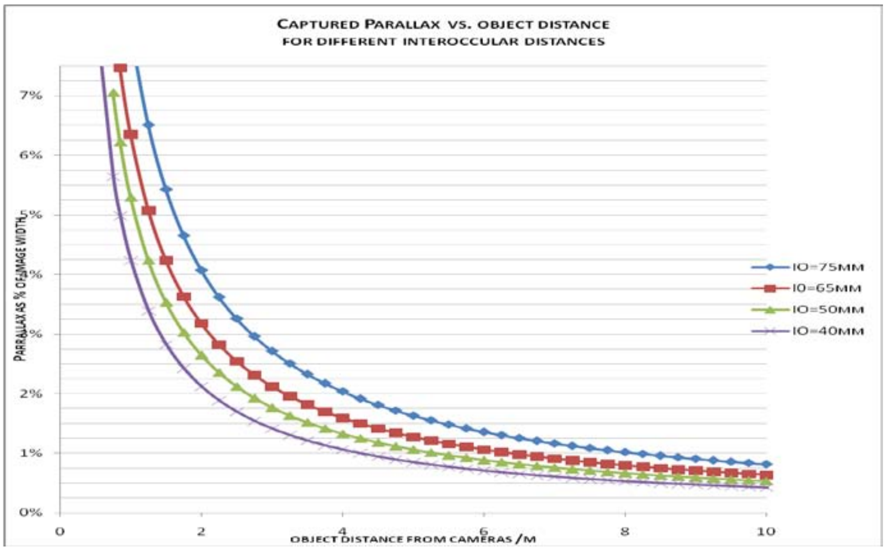
3.2.2 Camera Interaxial Distance (I-0)

In a 3D capture system the distance between the two cameras is called the ‘interaxial distance’ or ‘intra-occular’ distance, ‘stereo baseline,’ or simply ‘I-O’. In this section we consider how this affects the capture depth range, and user experience.

The average human intra-ocular separation is 65mm. Using an IO value greater than this is termed *hyperstereo*, while using a lower value is termed *hypostereo*. Excessive hyperstereo or hypostereo can give images an unnatural feel, known as miniaturization or gigantism, as captured depth is disproportionate to the relative distance between objects.

3.2.3 Camera I-O and Depth Range

The following chart shows how parallax, expressed as a percentage of the image width, varies with object distance from the camera pair.



This allows us to define a sweet spot as the 'near depth range' as being the range between 1% and 5% of parallax as a percentage of display width we get the following table. *The selection of these numbers is somewhat arbitrary, but is a useful exercise to illustrate the I-O depth range relationship.*

In broad terms, the 'near depth range' then effectively means:

- Objects within this range will easily be discerned as having different depths even on screens with modest parallax resolution.
- Objects at a greater depth may all appear at same 'far away' depth, though they may be discernable from objects at infinity.
- Closer objects may be difficult to view, *though in many cases Movidius' automatic convergence technology can yield excellent easily viewable images.*

IO /mm	Near Depth Range /m	
	Min	Max
75	1.50	8.00
65	1.05	6.25
50	1.00	5.25
40	0.85	4.25

Note that as the above analysis is based on parallax percentages rather than parallax angles or view angles, both of which require knowledge of display size and viewer distance. The assumptions used hold for typical handheld scenarios.

3.2.4 Summary and Key Recommendations

- A small amount of hypostereo gives a more useful depth range including close up objects.
- Too much hypostereo (i.e. <40mm) can give gigantism and a reduced overall depth range.
- A depth range of 50-60mm is a good balance, though lower may also be used.
- Good autoconvergence processing is required to maximize useful depth range.
- A wide horizontal view angle is useful for close up objects.

3.3 Rectification Processing

3.3.1 Overview

A captured pair stereo image or video sequence should have some differences due to different viewpoints which result in image parallax. However parallax should be the *only difference* between the two images.

Rectification is the process of correcting images so that they are identical in every respect, except for the desired parallax difference.

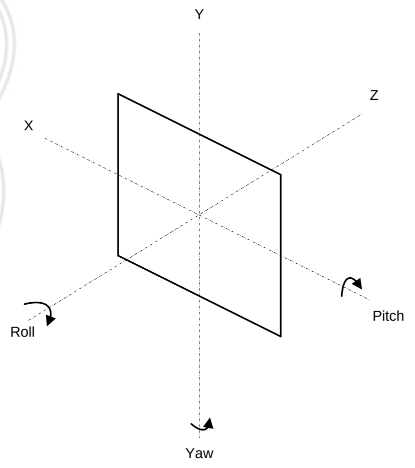
3.3.2 Camera Differences

Due to camera sensor and lens subsystem variation can introduce further sources of disparity between the stereo image pair.

Camera color balance and white balance introduce disparity in terms of pixels having slightly different values in both images. This can occur due to subtle differences in individual sensor gains.

Focal length and Field of View (FOV) variances in lens/sensor subsystem can introduce subtle differences in the relative zoom of an image pair.

3.3.3 Sensor Mounting Differences



Error	Meaning	Effect on image
Err_X	X translational error	Very small difference - effectively this is tolerance on the I-O chosen
Err_Y	Y translational error	Will introduce a very small depth dependant difference
Err_Z	Z translational error	Very small difference in image area
Err_Rx	Pitch error	Can cause significant vertical mismatch between images for small differences
Err_Ry	Yaw error	Can cause extra horizontal vergence/divergence beyond what would normally occur due to parallax
Err_Rz	Roll error	Above 0.5- 0.7 degrees of mismatch this becomes quite noticeable.

3.3.4 Rectification Requirements

Depending on the camera subsystems selected, a range of different corrections may be required to be performed actively in real-time during image/video capture. For system integrators there may be increased BOM costs associated with camera modules with tighter tolerances, so a 3D Capture solution employing a broad range of rectification capabilities yields significant benefits.

The rectification requirements based on the error sources already noted are:

- Detect/Correct vertical offsets
- Detect/Correct rotation mismatches
- Detect/Correct color imbalances
- Detect/Correct relative zoom errors
- Employ a convergence algorithm tolerant to horizontal offsets is required

Movidius' Myriad3D capture solution supports each of these rectification requirements.

3.3.5 Rectification Requirements for Asymmetric Sensors

A system where both camera modules are not the same model number may be termed Asymmetric 3D capture. In this case, rather than using 2x5MP sensors, perhaps an 8MP and a 2MP sensor could be used.

Motivation for such a system configuration is twofold:

- BOM cost savings through using lower resolution second sensor.
- Second sensor can have a higher pixel geometry, and can be used for 2D video capture. This can yield better quality 2D video capture performance, especially in low light.

Asymmetric 3D does introduce some extra challenges:

- Sensors will have a wider manufacturing tolerance (in symmetric processing it may be reasonable to assume sensors are from same manufacturing run).
- Sensors will have inherent differences such as view angle, resolution, pixel geometry.

These extra requirements mean that very high quality rectification processing is required to give images with low disparity. In particular:

- Color imbalances will be exaggerated, and require careful processing.
- Sensors resolution and Fields of View must be matched. Automatic detection of relative zoom errors is helpful here to allow for wider tolerances.

Movidius is developing the capability to support Asymmetric Sensors on the powerful Myriad 3D platform. Stay tuned to www.movidius.com for details.

3.3.6 Rectification Calibration

Two options exist here

- Specific static calibration step at assembly time. This can increase manufacturing costs.
- Automatic Calibration allowing cost savings. Allows for fine tuning of values over lifetime of handheld device to account for differences due to environmental factors (heat, cold, being dropped etc.).
- Movidius Myriad 3D solution supports Automatic Calibration.

3.3.7 Area Loss and Recovery

Note that all rectification (and convergence) processing results in a loss of area. Several approaches are possible here to recover the original area:

- Run sensor at higher resolution than captured resolution.
- Have a fixed black border around content.
- Use a high quality zoom algorithm to recover area.

The overall area loss will depend on final system configuration and tolerances. In particular even small rotation corrections can significantly reduce the useful area. Typical total rectification loss, including rotation, can be ~5% of the image area.
Maximum convergence area loss will be in the range of 1.5% of image width. Maintaining aspect ratio gives a convergence area loss of 2.10%.

3.3.8 Recommendations

- A fully featured Rectification algorithm can improve BOM cost and improve quality.
- Area loss is an unavoidable side effect of rectification and convergence.

3.4 Convergence Processing

Convergence processing is a process which adjusts the 3D in a captured image to give an impactful, yet easy to view result.

When a 3D photograph is taken, typically all content is captured with negative parallax. Without adjustment this will be unsuitable for viewing, and the photos must be converged for comfortable viewing. Convergence processing entails the balance of positive and negative parallax within a scene for optimal viewing.

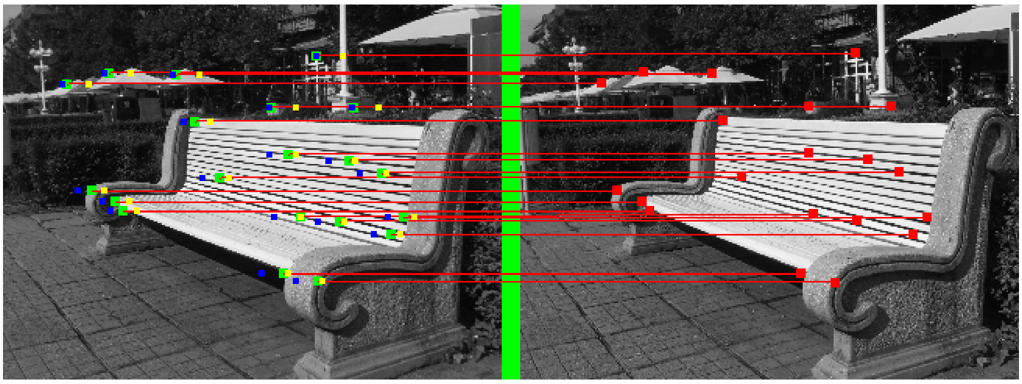
3.4.1 Fixed Convergence

One approach to convergence to is to use a smaller I-O distance between the cameras and use a fixed convergence setting. This is simple to do, and provides reasonable 3D effect for objects close to the camera. The downside is that any object beyond a fixed short distance (typically ~2m) from the camera will be captured with positive parallax and appear behind the screen. For a scene with mainly mid - long range content, all content will appear deep inside the screen.

Cameras and camcorders with small I-O distances therefore avoid the need to provide quality automatic convergence algorithms at the expense of an unreal and far-off 3D effect.

3.4.2 Automatic Convergence

On a platform with Automatic convergence, feature based analysis is done on the scene content to determine the appropriate convergence point. This allows larger I-O distances to be used, which give a more natural feel to a scene, and also provide a greater depth range.



For example, capturing a scene where the closest object is 5m away; this object can be placed slightly in front of the screen, and other objects a little behind the screen, enabling an immersive 3D experience.

A good auto convergence algorithm should also be able to provide some preset levels such as the following:

- Positive — Keep as much content as possible behind the screen, i.e. all positive parallax.
- Balanced — Capture 70% of content with positive parallax, 30% with negative parallax.
- Max3D — Capture as much content as possible in front of screen, within safety limits.

3.4.3 Automatic Convergence and Video Capture

Convergence processing is relevant for both video and still image capture in the 3D system. For video capture, the convergence algorithm should be real-time, and also must take care to adjust convergence gradually if required, to avoid any user discomfort.

3.4.4 Summary and Key Recommendations

- Camera axis should be parallel.
- Cameras must be vertically aligned.
- Larger camera FOV is best.

3.5 3D Storage and Transmission Formats

3.5.1 Overview

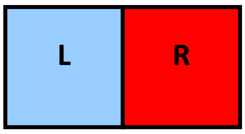
Many 3D standards for encoding and transmission have come and gone in recent times. In 2010 some practical clarity has emerged, and some key standards have gained traction among handheld platforms.

Standard	Description
H.264 AVC with Frame Packing SEI [1]	Various schemes, such as side by side (half) are possible for encoding stereo video within a full 2D video frame. In March 2010 a Frame packing SEI (Supplemental enhancement Information) extension was introduced to H.264 standard allowing the particular stereo packing scheme employed to be made available at playback time. This facilitates the development of systems that can offer seamless support of both 2D and 3D video content
H.264 MVC [2]	This multiview coding extension to H.264 allows for multiple views to be encoded within a single compressed bitstream. Key advantages of this approach are that it is compatible with 2D players and offers efficient full resolution encoding (typically 130% that of comparable full resolution 2D content).
HDMI 1.4 [3]	In Early 2010 the HDMI consortium publicly released a draft version of their 3D standard formats. This standard is independent of encoding formats, but specifically mandates transmission formats over HDMI physical interface. Full resolution frame packing and side by side (half) formats are supported as key formats, along with a number of others

3.5.2 Frame Format Examples

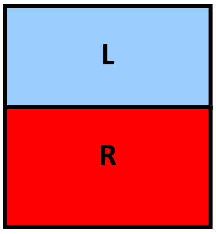
3.5.2.1 Side by Side (Half)

Side by side frame format gives results, but introduces a loss of half the horizontal resolution. A key advantage of this approach is compatibility with existing infrastructure.

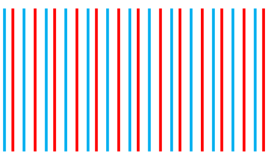


3.5.2.2 Frame Packed Full Resolution

Frame packed image format allows both images to be stored at full resolution.



3.5.2.3 Vertical Scanline Interleave



Vertical scanline interleave is common for use when driving parallax displays. One drawback of this format is that it is not compatible with chroma subsampling as employed in YUV420 or YUV422 video formats. This means that while commonly used in parallax display interfaces, it is rarely used in transmission or encoding.

3.5.3 Still Image Formats

For still imaging the Multi-Picture MPO image format (.mpo) has gained significant traction in recent years. It's worth noting that a stereo JPEG format also exists (.jps), which was popular in the past.

3.5.4 3D Sharing

The Movidius 3D platform enables 3D user generated content. The side by side (half) video format supported is also compatible with YouTube 3D video support.

3.5.5 Summary & Key Recommendations

Movidius recommends that side by side (half) is the best 3D format for User Generated Video Content. This has a number of key advantages:

- Reuses existing codec infrastructure for capture/playback.
- Same bandwidth as 2D signal.
- Supported format for 3D Video upload to YouTube.
- Supported by HDMI 1.4a.
- Popular format (used by Fuji Fine Pix, Cable-TV).
- No resolution loss when used with handheld parallax barrier displays.
- This approach is technically as good as, or superior to competing half resolution 3D formats [4].

In time H.264 MVC will gain traction, motivated by full resolution display support and bandwidth efficiency. A programmable and flexible platform like Myriad 3D can quickly adapt to changing standards without changing silicon, while always maintaining user 3D quality.

3.6 2D to 3D Conversion

3.6.1 Overview

While 3D technology is gaining momentum, there is significantly more content available in 2D formats than 3D at the moment. This is likely to be the case for a number of years, which means that 2D to 3D conversion is a valuable feature.

2D to 3D conversion enables any 2D video to be converted to a stereoscopic 3D which may be viewed on a stereoscopic display. For handheld devices real-time conversion is important - content is converted as it is viewed.

3.6.2 Summary and Key Recommendations

- For the next number of years 2D content will be more prevalent.
- Realtime 2D->3D is a key feature for handheld platforms.

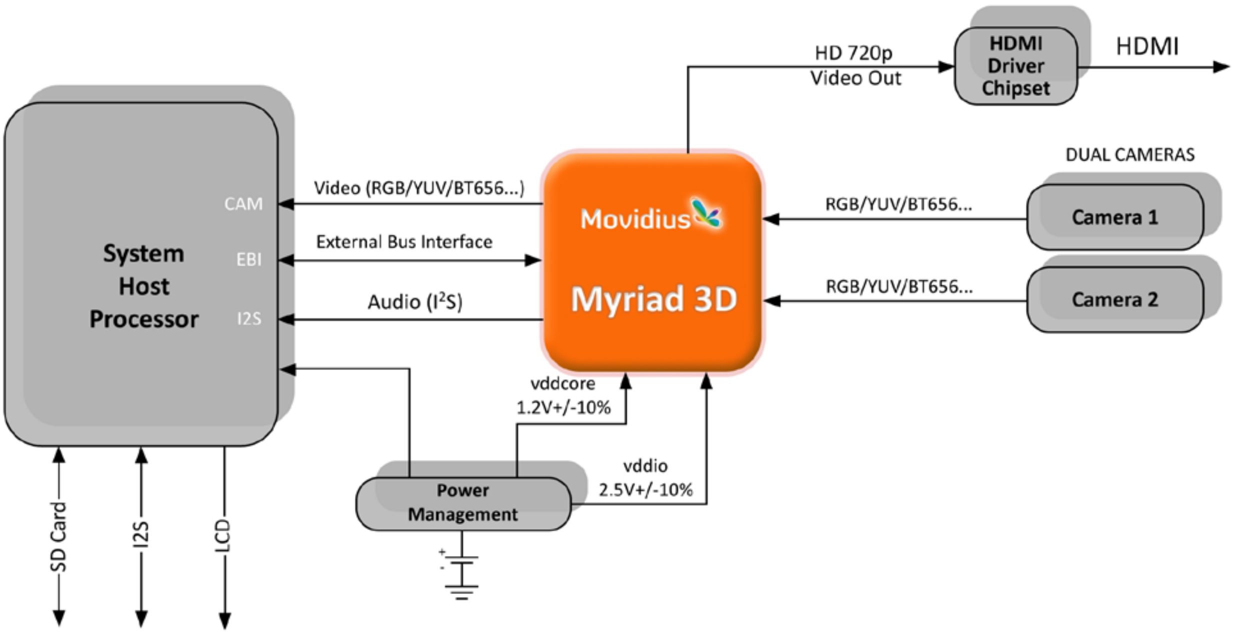
Myriad 3D offers a real-time 2D->3D conversion feature analyzing the content of each incoming frame to create two output frames for 3D viewing. The algorithm used ensures that a converted video is comfortable to view while also having an impressive impact.

4 Conclusions

4.1 Movidius Solution

Movidius brings deep expertise of video processing to the mobile 3D table. Movidius' focus has always been on the end user experience. The Myriad SoC platform is already sampling to customers bringing advanced video technology inside your phone.

The Myriad 3D family extends this platform's feature sets to bring a quality 3D viewing experience that is absent from so many early mobile 3D efforts to date.



4.2 The Total Solution

A good starting point = A Fantastic Finish

Just like a sprinter out of the starting blocks, a strong and powerful start to the 3D race means a comfortable and fast finish at the end of the race.

Showing 3D content on poorly constructed mechanics, adding an inferior 3D algorithm and constantly battling with identifying and removing gross design inaccuracies is a tough way to navigate the 3D space and the result will show (so will the headaches).

Building a tight relationship between design, manufacture and algorithms using the right components and technology gives the best starting platform for a successful and impressive 3D product. Add Movidius' rectification and convergence enhancements and allow the system BOM to be minimized by relaxing the mechanical tolerances through software compensation.

Enable the generation of 3D content with a quality HD 3D capture system, and in the meantime playback 2D movies as 3D in real-time and the ecosystem begins to form.

The 3D volcano is erupting and global mobile users will be there to witness Myriad 3D at the source.

5 References

[1] ITU-T-H.264_201003

[2] H.264 MVC Specification

[3] 'High-Definition Multimedia Interface Specification Version 1.4a Extraction of 3D Signaling Portion', March 2010

[4] 'Dolby Open Specification for Frame-Compatible 3D Systems', Dolby Laboratories, June 2010

[5] 'Foundations of the stereoscopic Cinema', Lenny Lipton, 1982

[6] ' Multipicture Format - CIPA DC-007-2009', Camera and Imaging Products Association (CIPA), 2009

[7] CNET.com

About Movidius

Movidius is a fabless semiconductor company whose technology delivers unique multimedia capabilities including High Definition 3D Video. Typical user applications include creating, editing, sharing, viewing and real-time improvement of multimedia content. Movidius Myriad and Myriad 3D technologies are specifically designed for low power mobile phone and consumer electronics, enabling manufacturers to create highly differentiated products and enhanced user experiences.

Movidius has offices in Dublin and Hong Kong, with a substantial software development center in Romania. Movidius is venture backed, with investors including: Celtic House Venture Partners, Capital-E, Emertec Gestion, AIB Seed Capital Fund and Enterprise Ireland.

movidius.com

youtube.com/movidius

Twitter @Movidius

Feedback

To contribute feedback, please contact Movidius Ltd. Include the following information:

- Myriad 3D, 3D White Paper, Movidius Mobile 3D - White Paper - 30101D