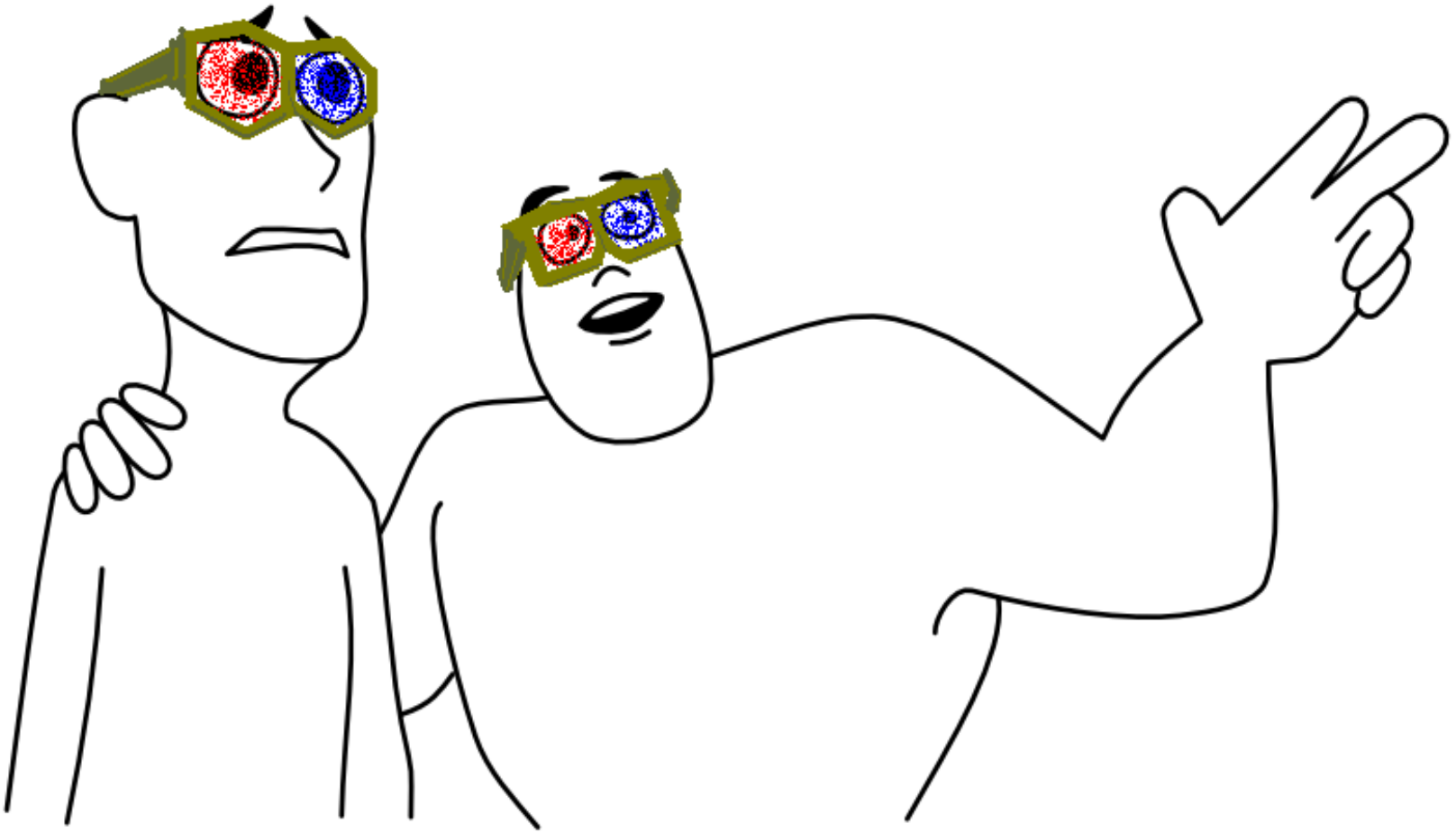
3D Imagery
Acquisition & Visualization

Marco Callieri
17/5/2012

3D Stereo



3D Stereo everywhere

The 3D hype...

3D, 3D 3D.... everything, nowadays, seems to have the need of the third dimension.. Movies, TVs, games and cameras/videocameras

Unfortunately, this trend will carry on for a while...

Important to know and understand the strengths and weaknesses of the available techniques...

The ability to generate and display stereo images is becoming a must in the visual industry... (and this is true also for cultural heritage applications)

Recurring

This 3D hype is nothing really new (very few things are really new in this world)... more or less every 15-20 year (beginning in the '50), there is a flourishing of 3D media...

Until now, each of those wave have faded quickly...

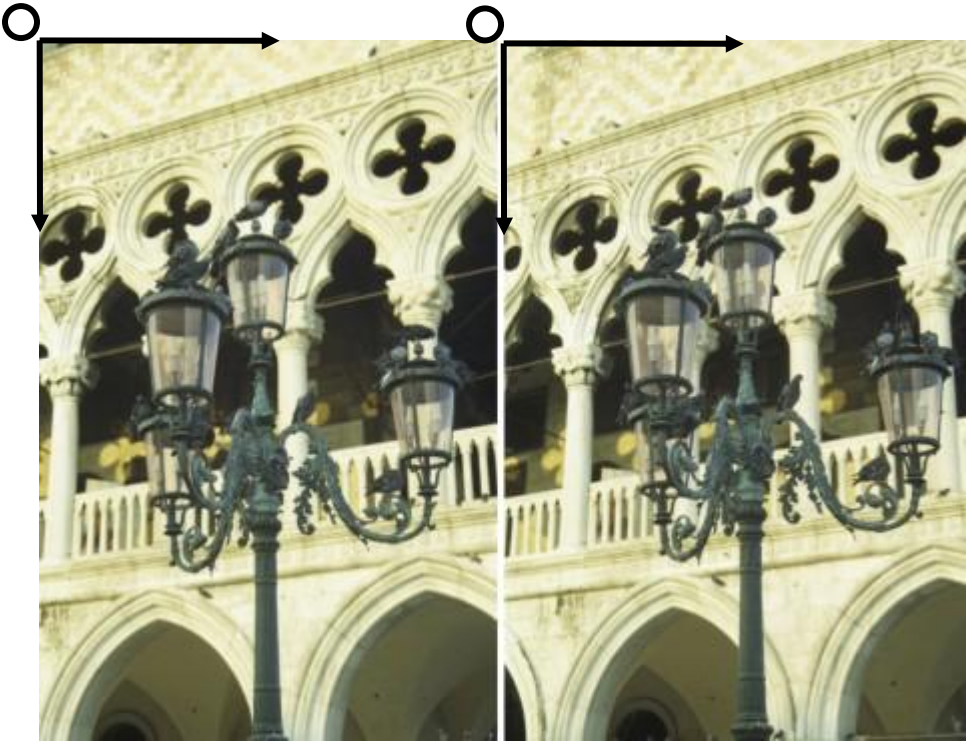
The problem is that now the technology is more mature, and the companies more willing to push a new product...

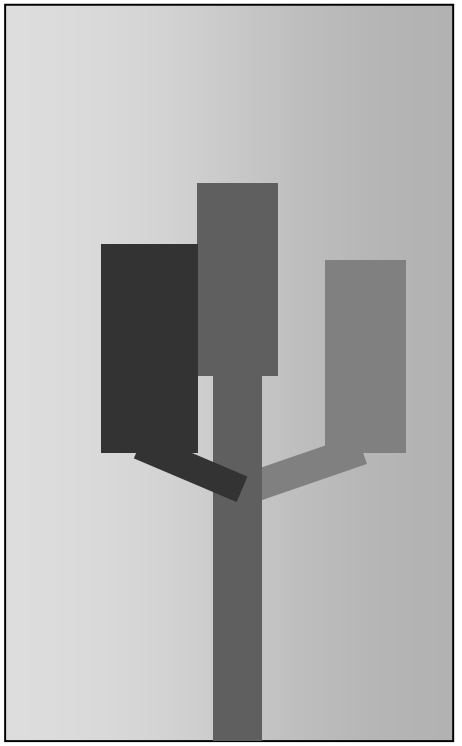
is 3D here for stay ? Hard to tell...

A simple principle

This is not about 3D data... but to exploit the 3D perception mechanism built-in in our brain.

All revolves around the acquisition/generation of two images and in finding a way to make each eye see the correct image... Generally it is called 3D, but *Stereo 3D* would be more accurate....



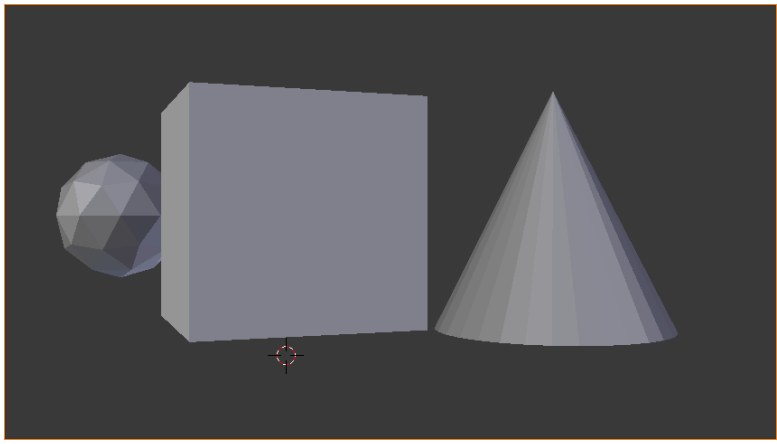


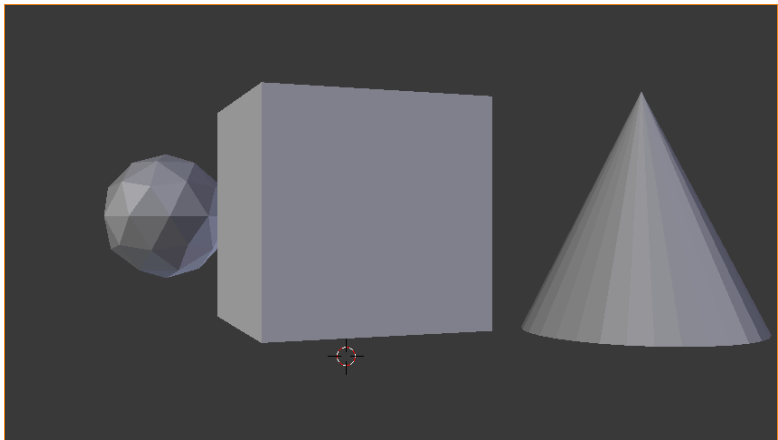
Some basics

Parallax: looking at a scene from two points of view makes things “move” from one image to the other... this *apparent movement* does depend on the distance of the object... The *amount* and *direction* of this movement (called DISPARITY) makes the brain understand spatial relationships between objects.

Here is a basic example: the cone has more offset than the cube, which has more offset than the sphere... With this, we can guess the order in space

(we also use other criteria, like the sphere occlusion, but without any reference it would be difficult to guess the position of the cone)





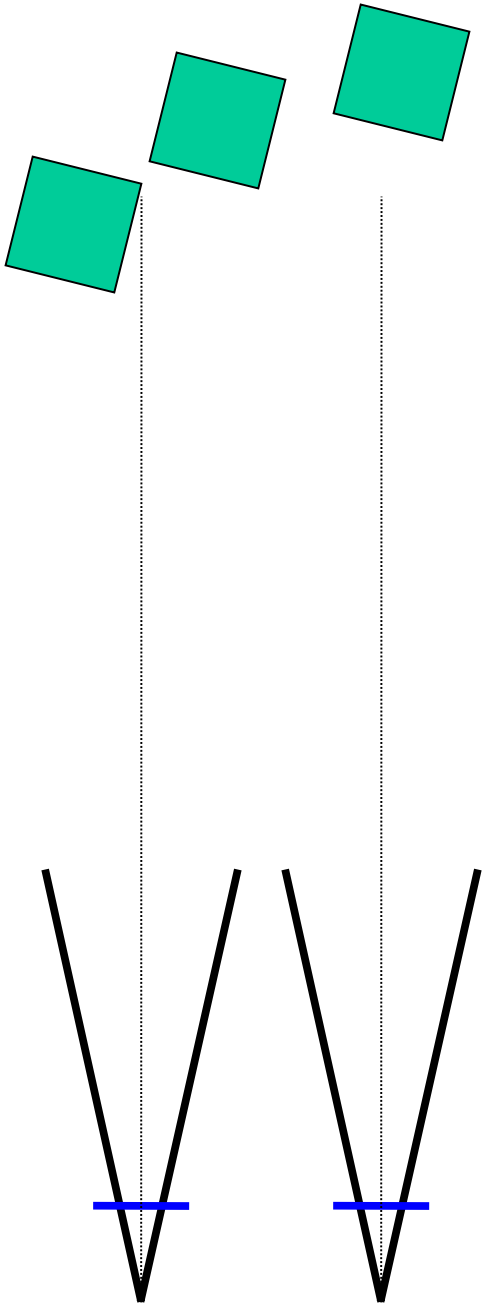
This is correct, but in reality it is not this simple...

The ideal model

The first idea of stereo couple is when we have two identical cameras, aligned, with an amount of horizontal offset between them...

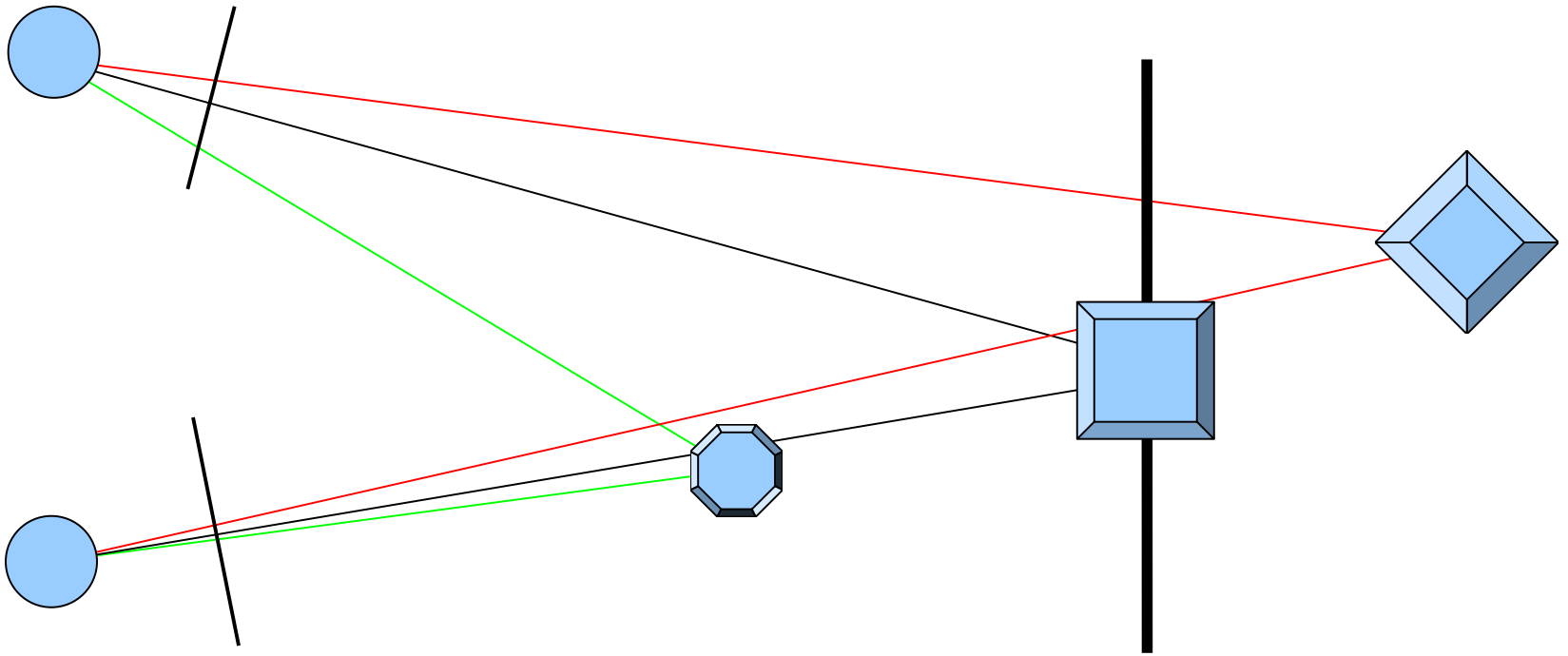
The distance between the cameras is called SEPARATION.

This model is, in principle, correct... but it is not exactly what happens in our eyes and brain... however, can be used without problems...



Positive and negative

Given that there is a “focal area” where our eyes are converging, that area will appear in the same place in both images. This, for our brain, is the center of the 3D space. Objects nearer than this place will have left eye image more on the right with respect to right eye image (*negative parallax*). Objects further away from this place will have left eye image more on the left with respect to right eye image (*positive parallax*)

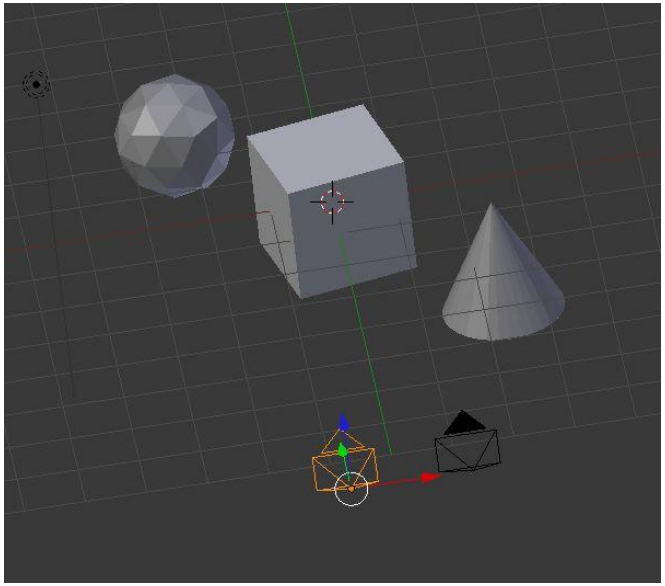


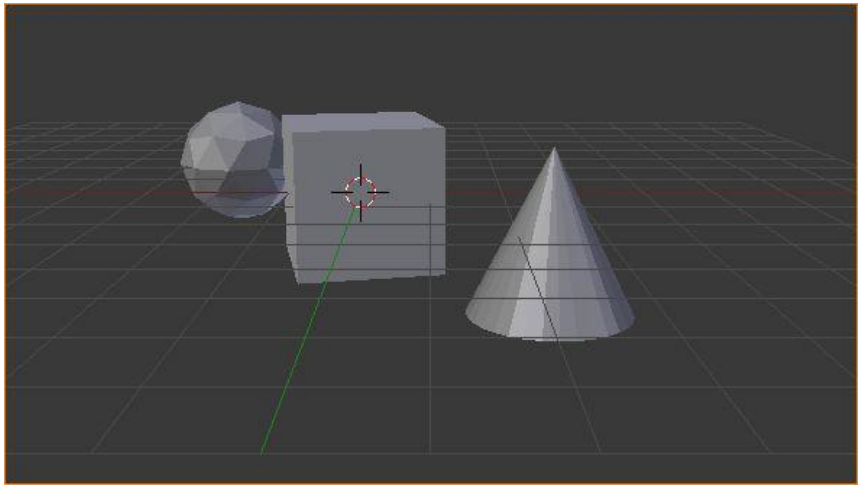
The simplified model

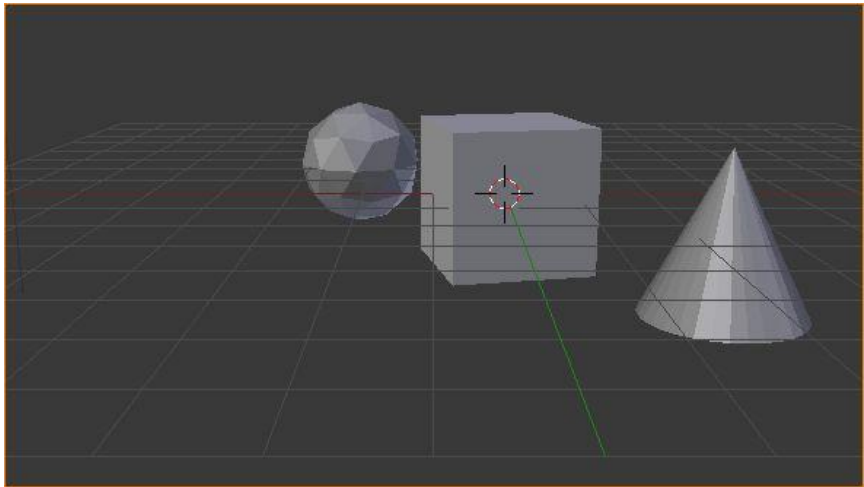
Keep the cameras straight, do a manual convergence on the images by shifting their overlap to make the “focal area” coincident. The focal area will be perceived at the same depth of the screen

In this way it is possible to *manually* select the focal area....

If the focal area makes sense, the user will perceive the scene much better

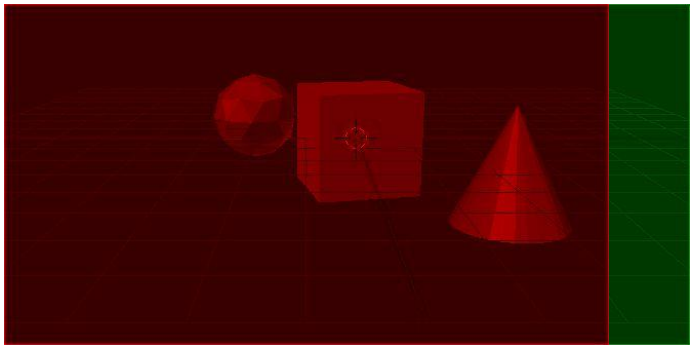






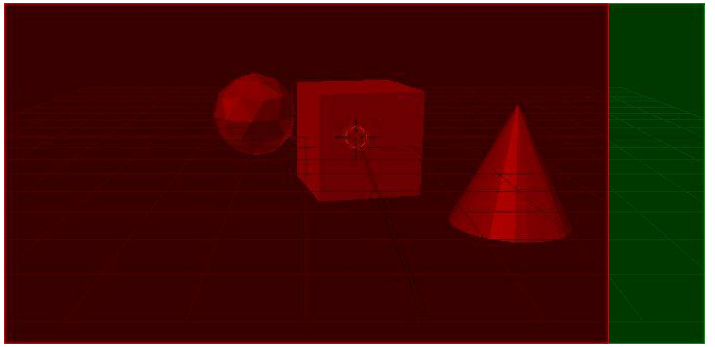
The simplified model

Aligned on sphere: sphere is on monitor plane, cube and cone in front

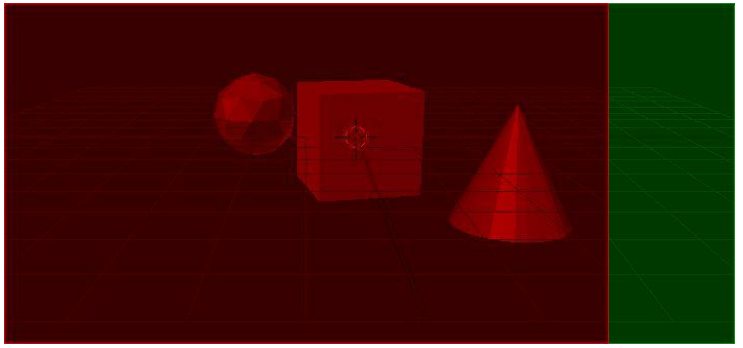


Aligned on cube: cube on monitor plane, sphere beyond and cone in front

(note how parallax is inverted between sphere and cone)



Aligned on cone: cone on monitor plane, cube ans sphere beyond



Spatial relationship between objects DO NOT change (sphere back, cube in the middle, cone in front).. it is simply a different focus area

So what ?

Regardless of the method used (ideal or simplified) we will find 3D stereo disorienting because we will NOT be able to focus our attention as we would normally do in real life...

To perceive the stereo effect, it is necessary to relax our eyes to accept the artificially-set focal point.. This, however, it is not easy, since the brain does it automatically: hence, we can experience a sense of uneasiness / dizziness.

The separation dilemma

We have a separation of around 10cm in our world. This mean that the 3D stereo effect in our brain works incredibly well under 1 meter, so and so between 1 and 3 meters and is practically useless beyond that...

To make larger scenes perceivable in 3D it is necessary to enlarge the separation... however, a completely out-of-scale separation will make the brain uncomfortable...

Outline

Given these premises, the problem can be split in its basic components...

Acquisition: how to acquire a stereo couple

Generation: how to render a stereo couple

Visualization: how to visualize a stereo couple

Miscellanea: why things does not always work

Acquisition

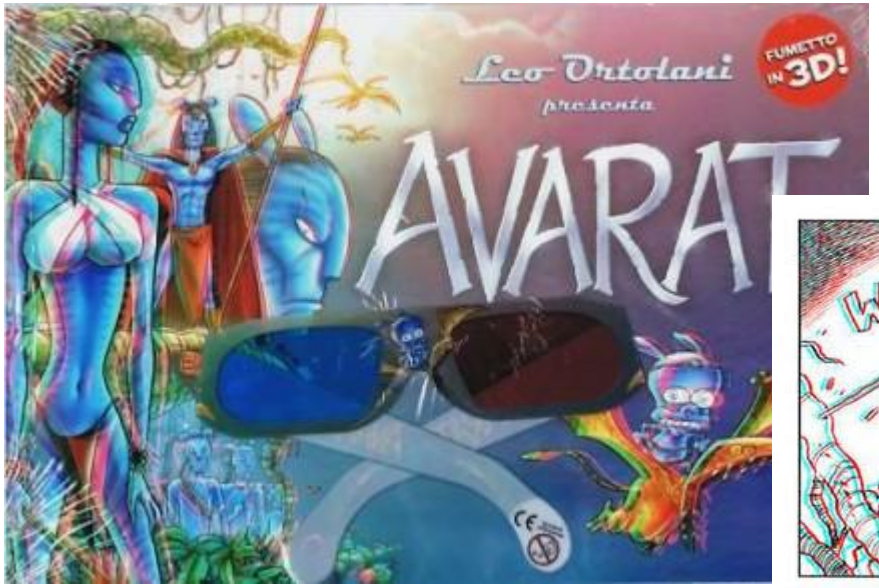
A genius...

Around 1971, Salvador Dalì created *hand-drawn* stereo couples...

Dali's Hand Drawing Back the Golden Fleece in the Form of a Cloud to Show Gala the Dawn,Completely Nude, Very, Very Far Away Behind the Sun

A crazy guy

Leo Ortolani, in 2010, published “AVARAT”... a 3D comic book

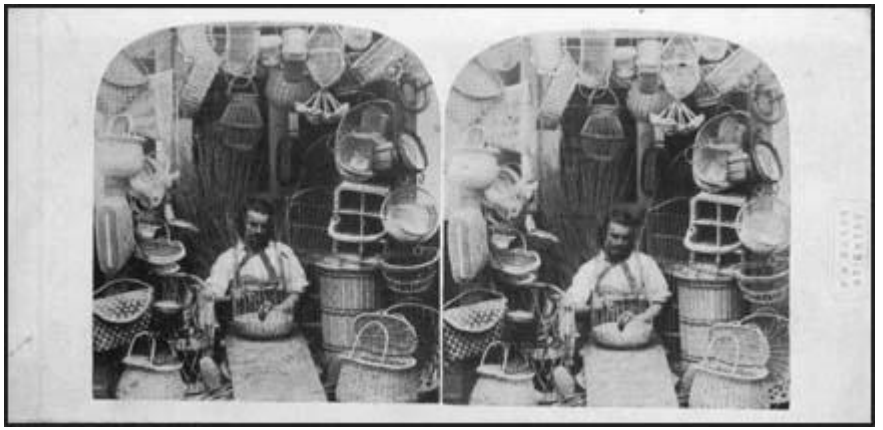


Actually, it is not the first 3D comic, but it is incredibly good at timing :)

Old tricks, new gears

It is important to say that stereo photography as almost as old as photography itself...

“stereo” cameras with dual optics have been around for years...



Simple Idea | Idea

If I have two optics, I can get a stereo couple !

Well... yes and no...

It is not easy to change separation and it is not possible to have a proper convergence (yet).

Additionally, lot of care should be put in taking both images with the same parameters (exposure, timing, focus....)

People had taken stereo photos by using a rig of two digital cameras since a long time... and shift-aligned the images more or less manually.

Dual Optic

One year and a half ago, Fuji released a 3D camera with two hardware-synchronized lenses... focus and exposure are linked and try optimizing both images. Alignment is pretty good and stable.

Images are saved as a couple (but in a proprietary format).

It is possible also to capture stereo videos...

This camera is a bit bulky and rough,but it is more or less a prototype... other brands are working on similar devices...

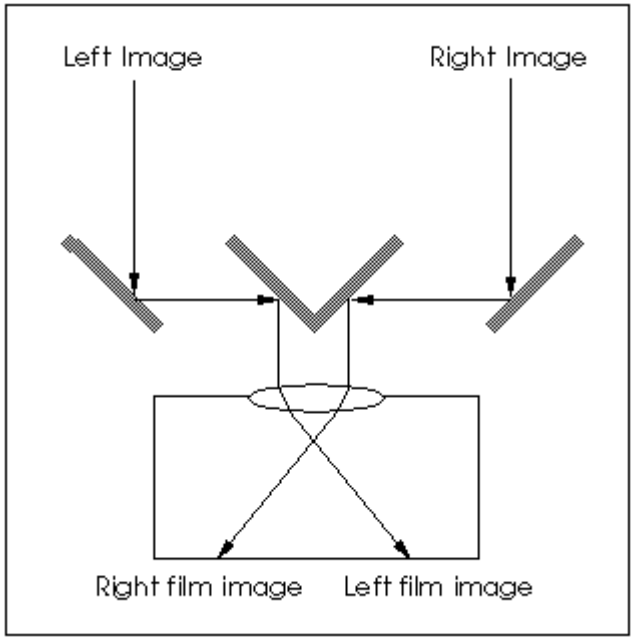
3D data can be extracted from the stereo couple using dense matching

Beam Splitter

With a single optic, I can get a stereo couple by using mirrors to create a couple of “periscopes”, each one generating an image on one half of the sensor/film.

(technically speaking, the term beam splitter is completely wrong, but in most commercial sites, these devices goes under this name)

- halves the resolution of the device
- distortion (mirrors are never *exactly* at 45°)
- problems in focusing and metering
- a processing tool is needed to separate the images accurately
- another lens in front of the camera



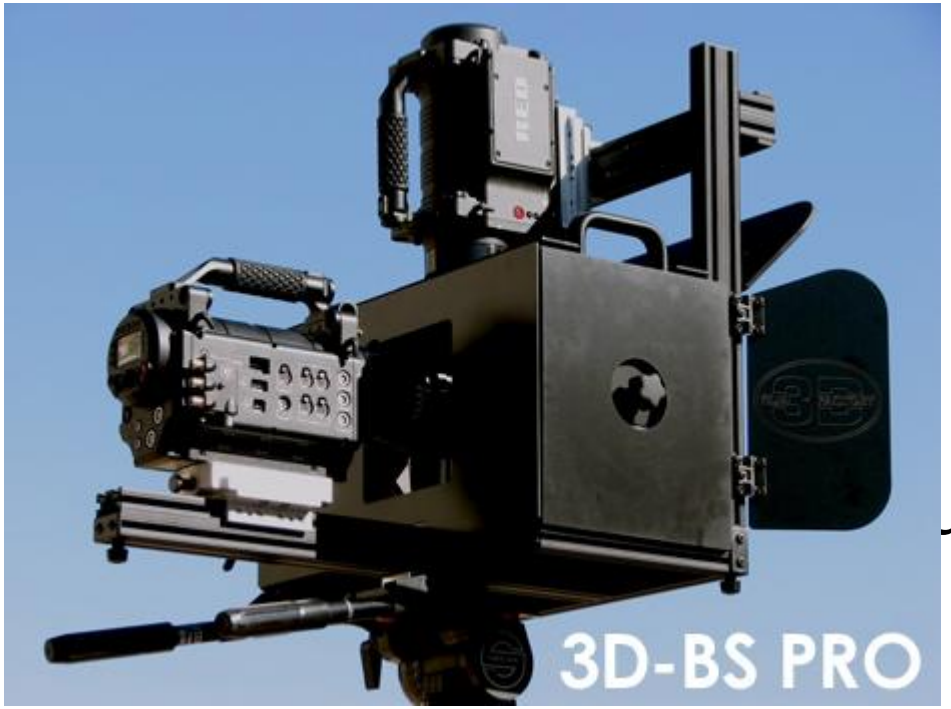
Beam Splitter

Various devices, from the “hobby” product able to fit on low-end cameras, to complex rigs for video cameras

Dual Cameras

For video acquisition, it is often used a rig with TWO cameras. Since most of the high-end cameras are digitally controlled, it is possible to sync focus/exposure and other parameters.

Complex mirror rigs (similar to the beam splitters just presented) are used to control camera *separation* and *convergence*



lots of effrt to contron the rig

lots of processing needed

However, this is what most people
uctions

Single Optic

With a single optic, I can get a stereo couple by processing two images taken from different positions, if they are “more or less” the *correct* position...

Image processing technique are able to correct the images in order to make them look like taken from the correct stereo couple viewpoints

WARNING: part of the process does warp the images in order to make them match better... this result in a 3D usable for viewing, but not good for data extraction.

Sony “sweep stereo”

Sony recently introduced some compact digital camera able to generate stereo by sweeping the camera in front of the object of interest (other manufacturers followed this idea).

The stereo couple may be viewed on the 2D display by tilting it back and forth, or sent to 3D TV....

More than one image can be captured (multiview).. but it is impossible to display it on standard stereo devices.

Alice

The movie Alice in Wonderland (the latest one)

Almost no “real” 3D acquisition... most of the scenes were shot with a single camera

This gave the impression of an animated “pop-up” book.. and also gave many spectators a good deal of headache.

Generation

Two, only costs twice...

Since I already have the 3D scene and a way to render it, putting two cameras and rendering it two times, will produce a stereo couple...

Again... this is both true *and* false...

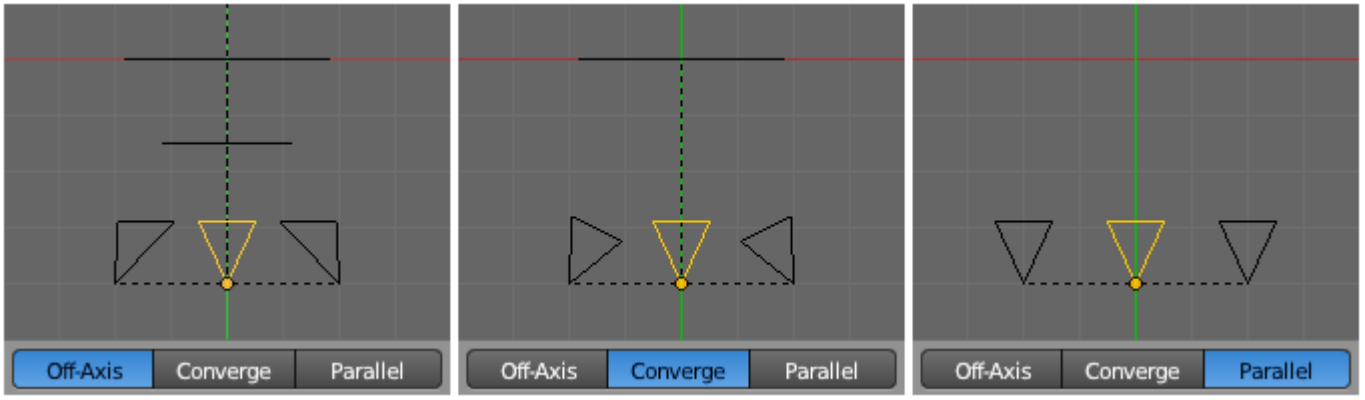
Putting two side-to-side, aligned cameras will have the same problems of the double optics systems... (separation and convergence)

It may be fine in realtime applications and when exploring single-object datasets, but it generally fail on more complex scenes (and in rendered movies)

Two only costs twice...

Especially for movie rendering it is possible to create “virtual rigs” of cameras, and control separation and convergence in a much easier way

For most of the 3D modeling&rendering tools are available such components, which control convergence using the image shift (as seen before), or real camera convergence



What cannot be done

Most of the post-processing effects used in image processing and movie production are intrinsically 2D and, more important, are just *simple approximations* (a.k.a. tricks).

Consider things like motion blur, defocusing, lens flares, fog, noise, glare...

Sadly, most of these effects does not work well on stereo frames... two solutions:

- compute the full 3D effect (much more costly)

- switch to something different

However, new possibility of post-processing are sprouting, which exploits the stereo pairs :)

Visualization

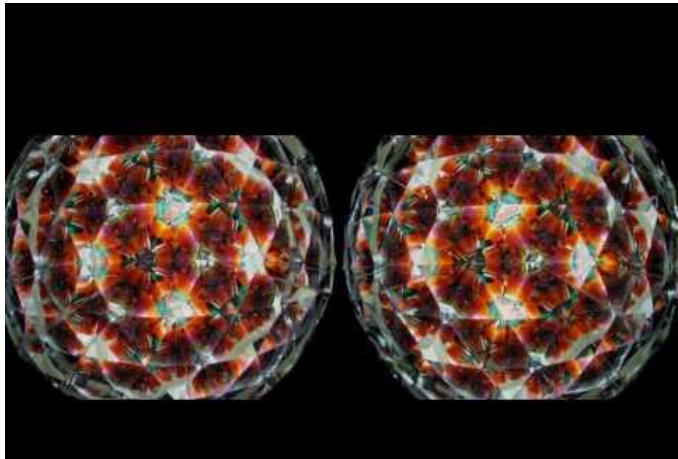
Side-to-Side stereo

A shockingly simple method is just to print the two images side by side and “defocus” your eye. At the correct distance, you will perceive the stereo effect...

Basically, the idea is to put the eyes in the position of the camera and focus them straight ahead at infinity (the way the image were taken).... small problem, you have to keep your face quite near to the screen

I say *you*, because I can't... I have a very dominant eye...

Believe it or not, many scientist use this method when looking at stereo couples...



Animated GIF

Another shockingly simple method is to create an animated **gif** which quickly alternates the stereo pair

Works really well when there is a “focal” object in the center of the image, where the zero-parallax is located...

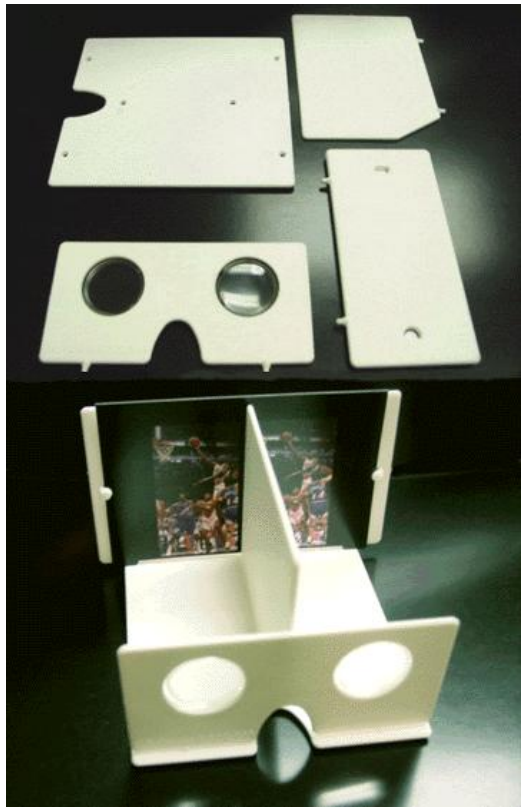
With some editors, it is easy to create these images, also tweaking the zero-parallax area..

Too bad they will not work inside this presentation :)

ViewMaster & friends

The two images are kept separate, and presented to the correct eye through a binocular-like device... The idea started at the dawn of photography, and evolved in more efficient forms later on

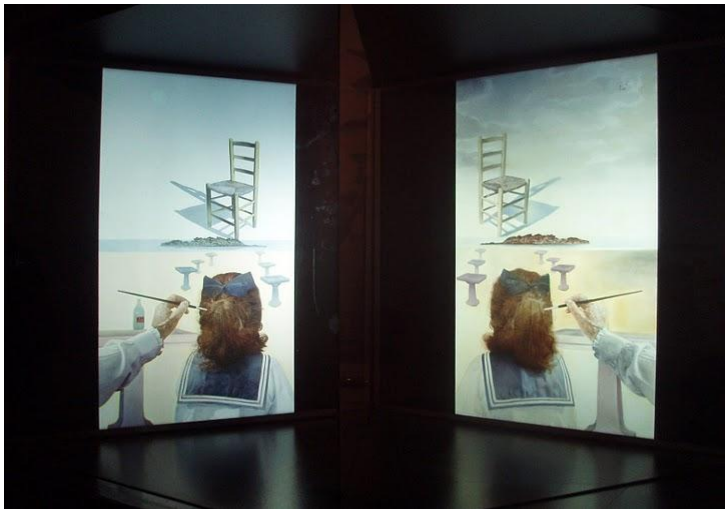
As an example, the ViewMaster, invented in 1938 and still produced as a toy/souvenir all over the world up to few years ago..



ViewMaster & friends

The 3D painting at the Dalì museum are displayed similarly..

There are devices which affixed to monitors and screens will do this simple trick...



Glasses

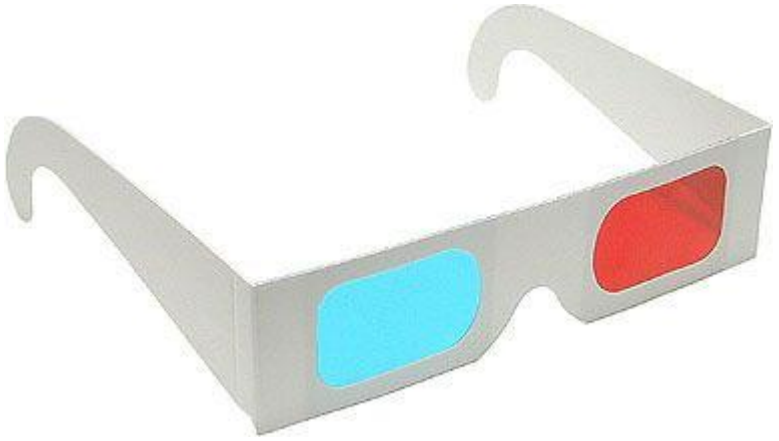
- Glasses-based methods are the most used... their cost-benefit ratio is the best one, even if they require the use of glasses
  (and if you, like more or less 50% of the world population, already have glasses, this is a bit unpractical)
- Mix the two images on a single plane using a specific strategy, and use some special glasses to separate them again just before entering the eye...

Anaglyph

The oldest and simplest method... build a composite image using the red channel of one side and the green+blue channel of the other side.

Using glasses with pure red and pure blue lenses will separate the images...

Every existing device (I would say, every media, paper included) can support this... however, high stress for eyes and brain, and severe color drift.



Anaglyph

Incredibly cheap...

Works really well for greyscale images...

The glasses may be red-blue or red-green... I always forgot which is the standard left/right configuration

More advanced techniques separate the color not in RGB but in other color spaces, thus reducing the color drift... the result are those pesky forest green - lilla glasses...

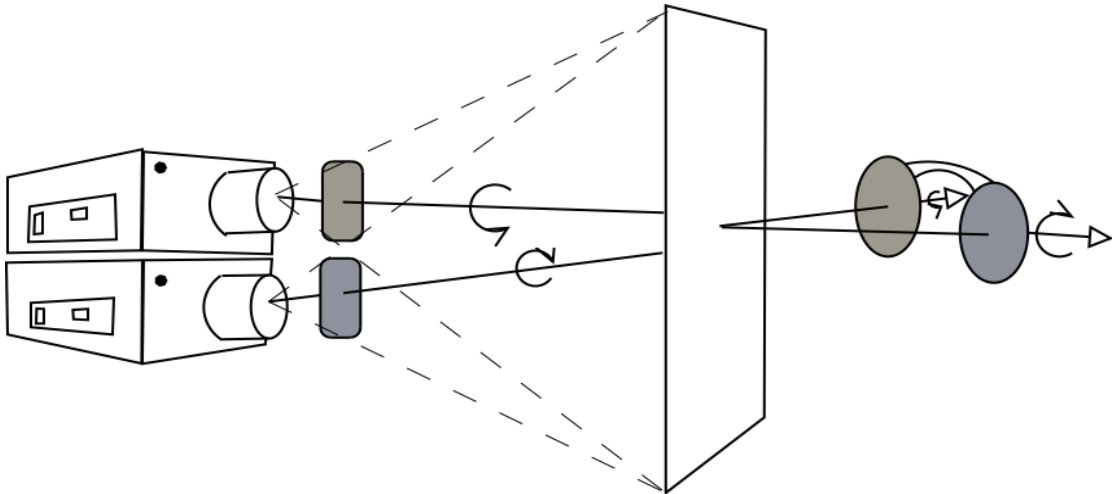
Polarized Anaglyph

Two projector, each one with a different polarization: the two images are at the same time on the screen, but still optically separable with glasses; each lens has the correct polarization...

Again, cheap but effective.... color is much better.

Most 3D cinemas uses this method...

- Require two projectors and a special screen
- It is necessary to stay in a narrow cone in front of the screen
- Saturation and brightness suffer a bit



Shutter Glasses

Images for the two eyes are displayed one at a time, alternatively (left right left right left rigth...)

The glasses, in sync with the display, shut the opposite eye. Synchronization was once obtained with a cable, now with infrared or radio link.

Many home 3D TV works in this way...

- It is necessary to have two times the refresh rate (120 hz)
- glasses need power and cost a lot...

Active Glasses

Different from the previous approaches: the images ere kept separated and displayed using two micro monitors, one for each eye...

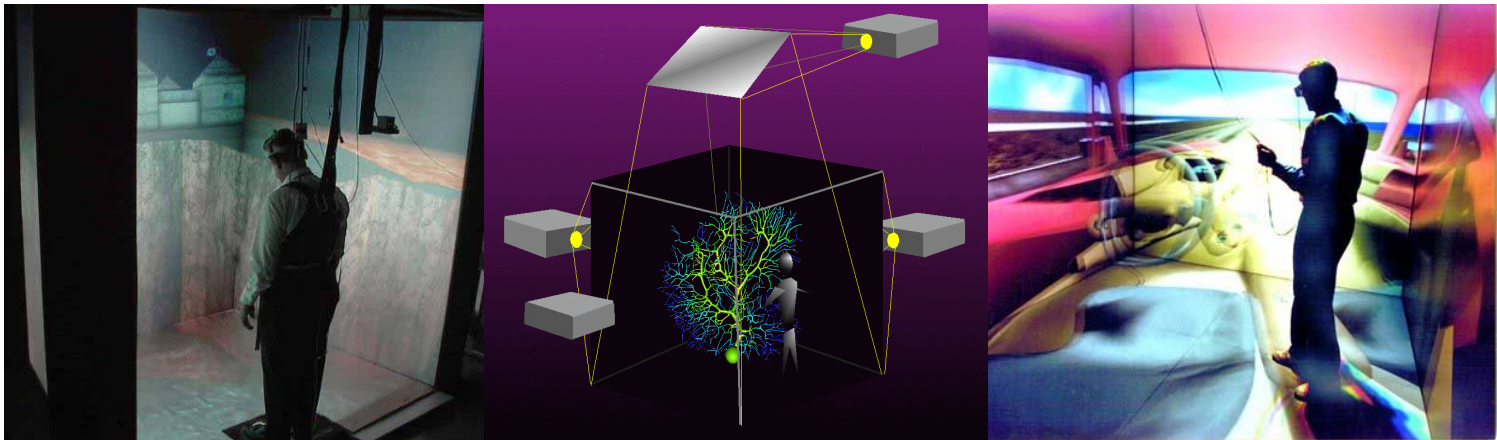
Extremely costly and not really standard... plus, the resolution is often quite low and having two retro-illuminated monitors at a couple cm from your eyes is not than comfortable...

Active Tracking

When the position of the user is known at each time (using some kind of tracking device), it is possible to generate stereo views on the fly following the user movement, which increases a lot the sense of immersion

This is possible using any of the glasses-based techniques

This was the idea behind CAVEs and Virtual Reality



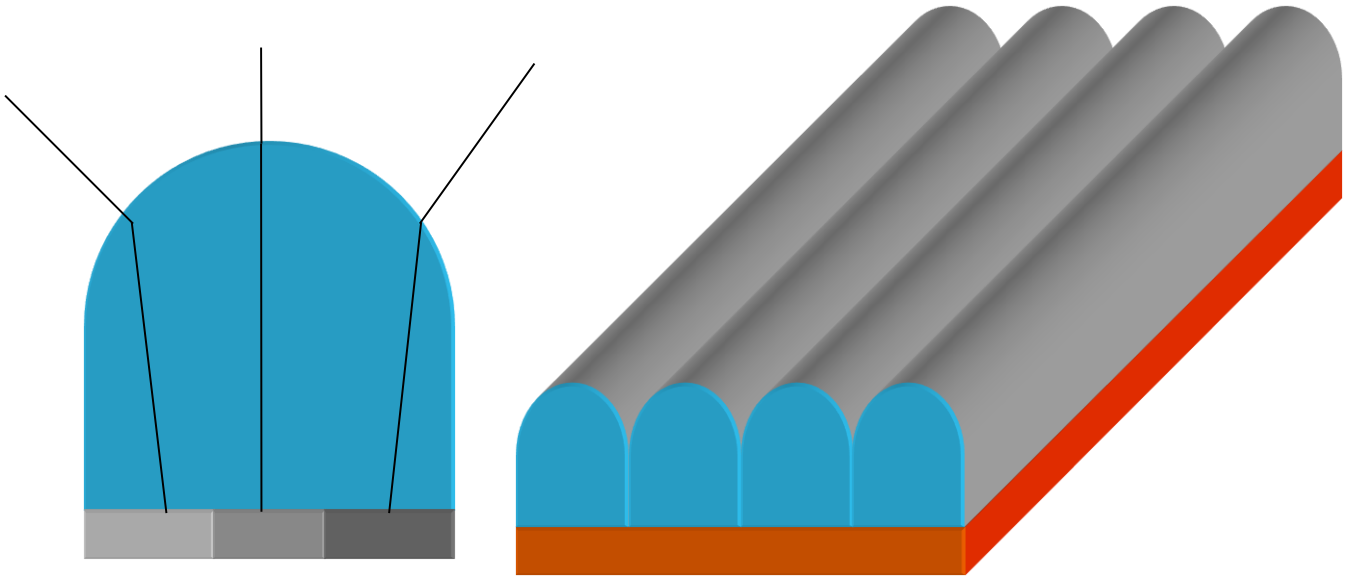
Micro-Lenses

It is possible to have 3D without glasses... for example, by putting the lenses on the display surface...

An array of micro lenses, able to convey different content when looked from different directions

The images are sliced and each slice alternatively printed side-to side...

If the dimensions are “good” each eye see a different image... and by tinting the image, it's like tilting the object



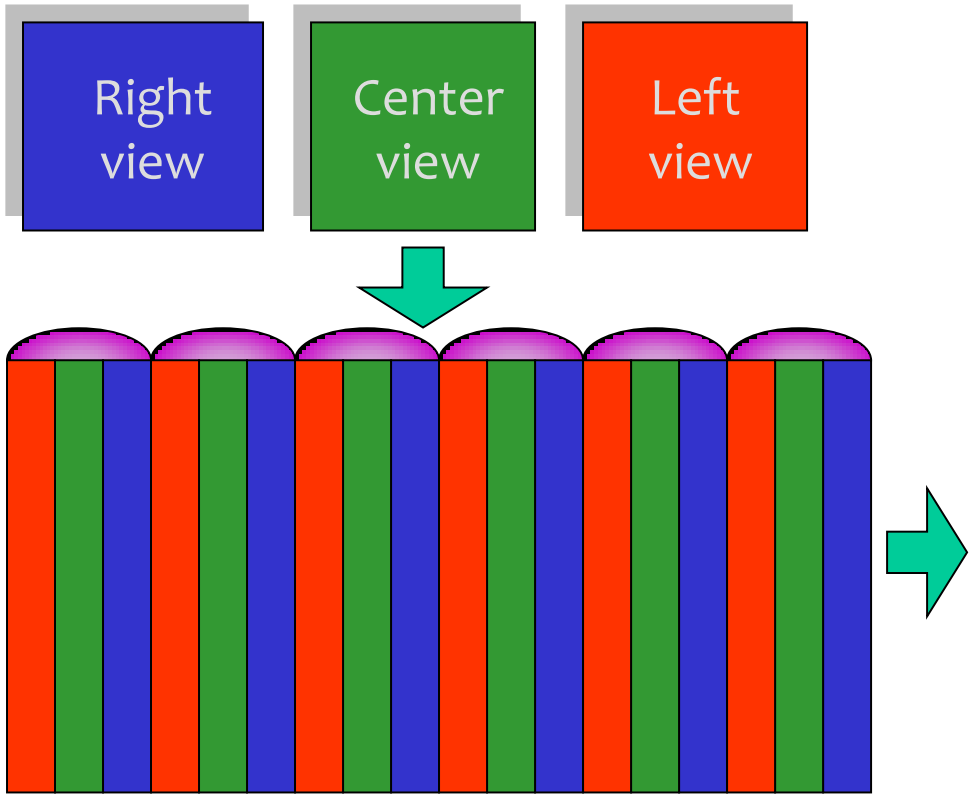
Micro-Lenses

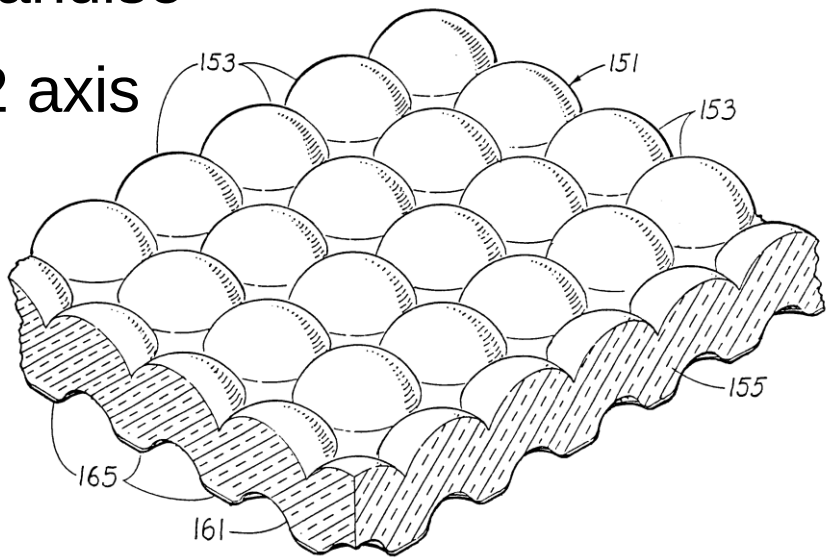
Many different images (>10) can be combined...

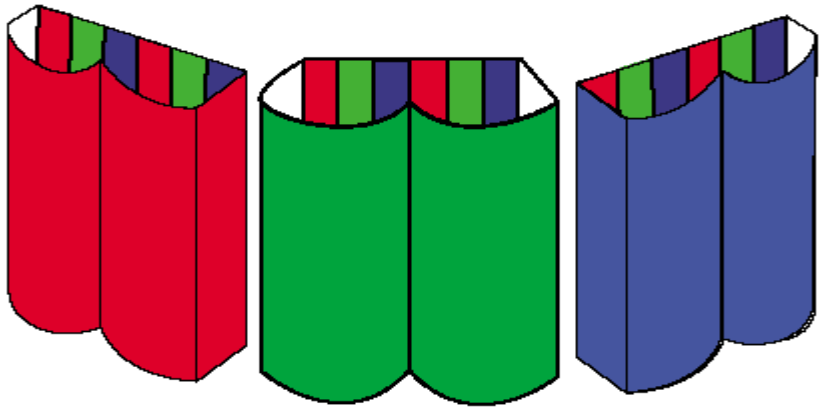
Easy to produce, even at home: a specialized software to mix the images and a sheet of lenses to glue on the printed page

Low cost but effective: highly used in merchandise

It is possible to have images with offset on 2 axis





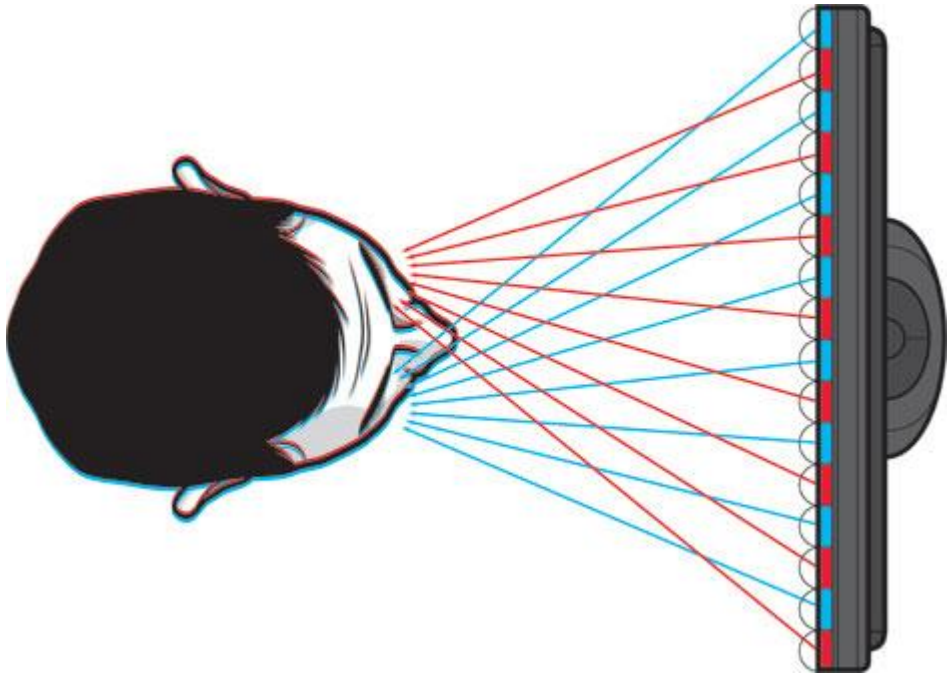


Lenticular monitors

Lenticular displays are available on the market...

Generally, restricted to two images, horizontally separated..

Costly, and with a very small “sweet spot”... However, suitable for small display (3D cameras, 3D camcorders, smartphones)

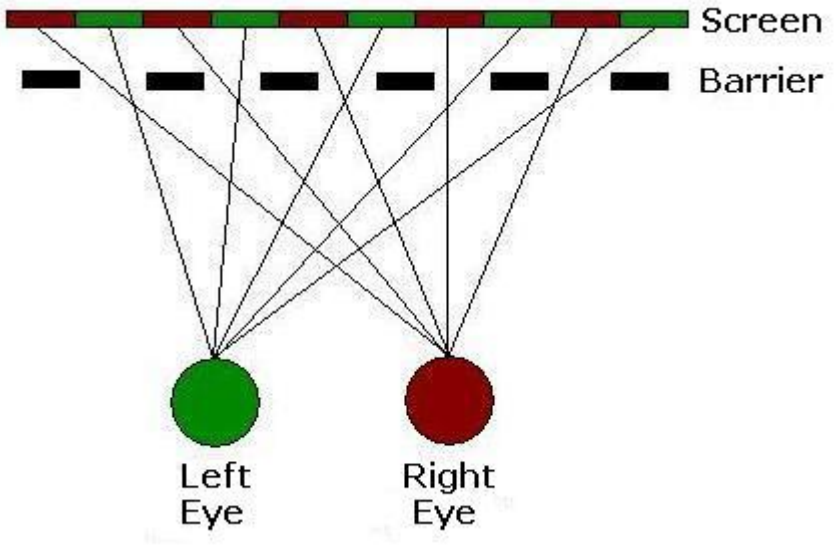


Mask-based

Similarly, it is possible to control which image is viewable from a certain angle through the use of a mask.

A “parallax barrier” composed by vertical bars, placed close to the screen makes visible a different part of the pixels depending on view angle...

The Nintendo 3DS works using this principle... some other mobile 3D display works similarly.



Mask-based

Similarly to lenticular technology, with a 2D grid instead of bars, it is possible to have 2 directions of parallax.

It is also possible (but tricky) to build your own 3D display modifying an LCD monitor (there was a tutorial at siggraph last year)

The interesting idea is that the barrier can be made using *another* LCD layer... which makes possible to control separation (and turn it off completely to revert to 2D image).

Miscellanea

Problems

3D seems incredibly fun on paper... but why lots of people does complain about it when viewing stereo images/video ?

The stereo technology try to piggyback our 3D vision system but, unable to address all the details of our vision mechanisms, may present incomplete or unstable input… this instability is normally associated in our body with NAUSEA

Additionally, every time the brain feel that something is wrong with the incoming perceptions, sends out a “warning” message... this gives the brain extra attention, which, in turns, bring us out of the suspension of disbelief :(

Measures

Each one has its own body measures... this also apply to eyes

- eye distance
- focusing range
- eye dominance
- ...

It is also true that our 3D perception (the perceptual part in our brain) is different from individual to individual so, from the stimulus, two brains will perceive a different scene

A single stereo couple will never be able to cope perfectly with the specific measures of everyone... It would be needed a system calibrated to our needs, and this is possible only in single-user applications...

3D is just 2D+

3D movies are just regular movies, but there are more things flying towards the audience (or spiky thing protruding from the screen)...

3D is not effectively used as an expressive tool, like the rest of video/image effects....

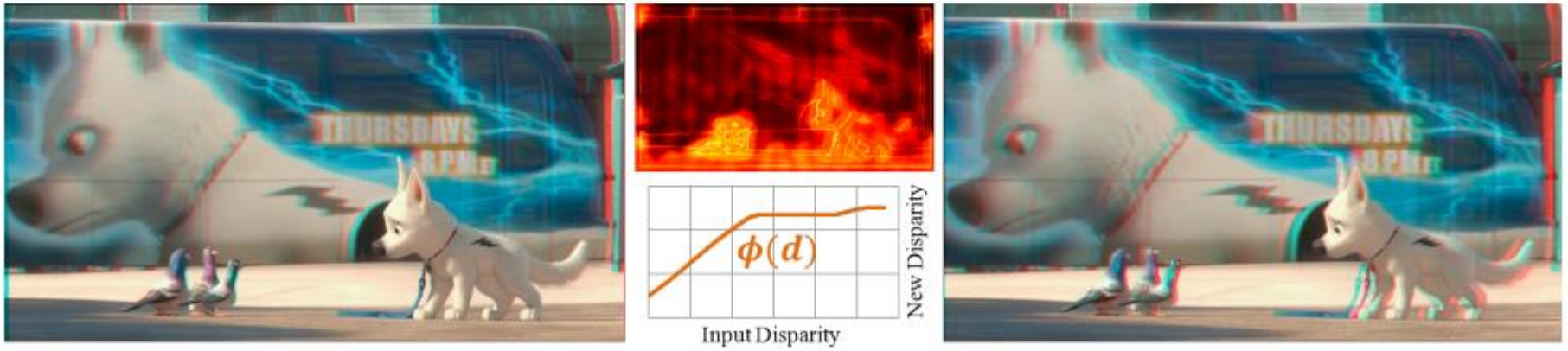
This is only partially a technological problem... the technical difficulties are a cause, but the lazyness of directors and the scarce innovation of the mainstream movie industry is much more responsible of this...

I have to admit that Cameron DID tried hard in Avatar... but he could have gone further...

3D effect editor

A paper at the last siggraph, by Disney research, presenting some image processing techniques for editing the 3D effect on stereo videos....

- smooth out sudden depth jumps in the scenes
- accentuate the 3D effect temporary, as an expressive tool
- slightly correct the extent of depth of field
- make the 3D effect less stressful for the eye



3D effect editor

The interesting part is that it operates on the video couples, by means of local warping/shifting of the frame couples, driven by image analysis tools and following perceptual studies...

This mean less work to do at capture time, more freedom to use the 3D really as an expressive media tool...

The methods are quite easy to implement... it may be possible for the next generation of video processing tools to have integrated such technologies...

To conclude

Some remarks:

- 3D as it is has not much sense, beyond being a toy.
  Unfortunately, it is a mandatory toy in today's entertainment.
  Technology able to control 3D effect and the possibility to use it as an expressive technique may open up new possibilities

- stereo sound seems more simple than 3D imagery, and in fact it is (we have less brain devoted to it)... however, it is not *that* simple either

- to prevent the classical question, YES, 3D porn does exist (remember rule 34)