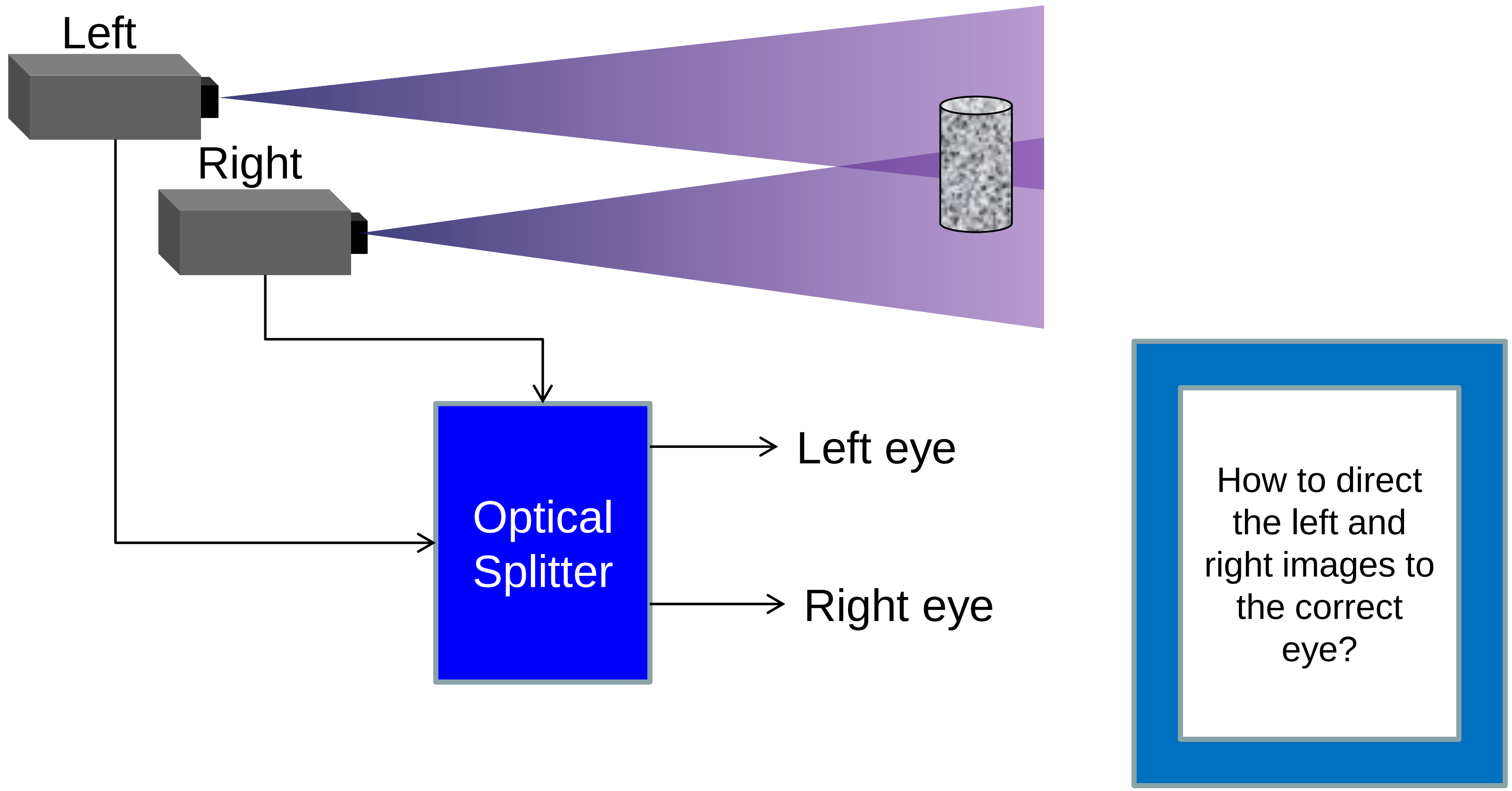
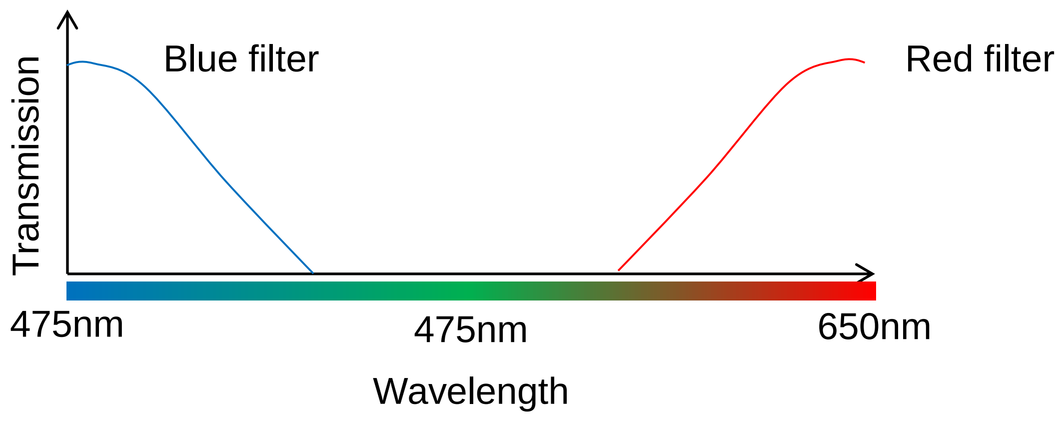
3D for the public viewers

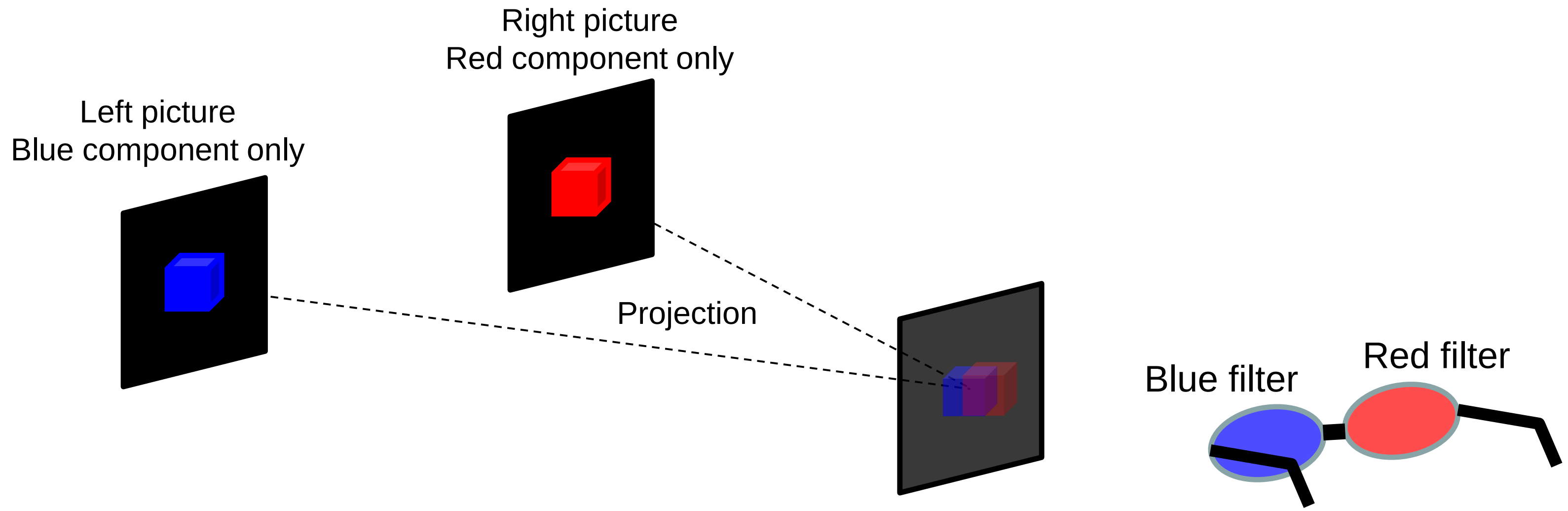
Binocular (perceivable with 2 eyes)



Anaglyph



Transmission
Blue filter
Red filter
475nm
475nm
650nm
Wavelength



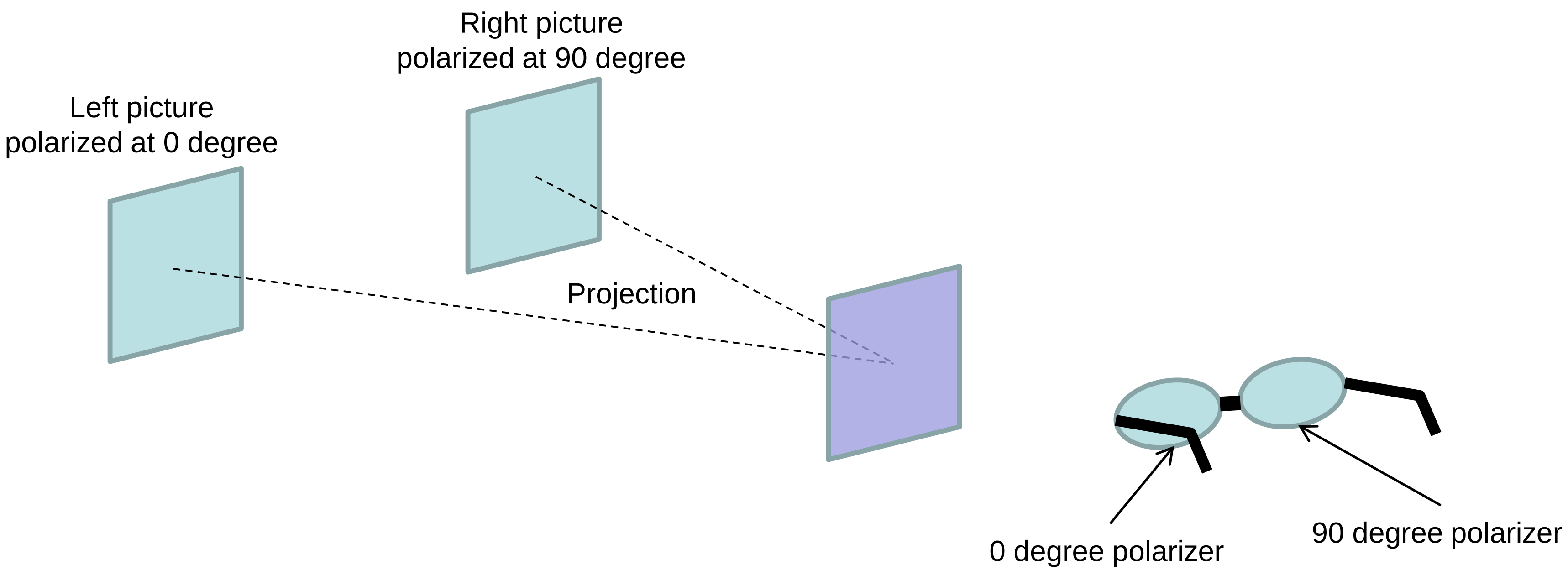
Right picture
Red component only
Left picture
Blue component only
Projection
Blue filter
Red filter

Polarized glasses based Stereoscopic 3D

Usual approach: Display left and right eyes images on the same screen, and utilizes polarized glasses to direct each image to the corresponding eye.

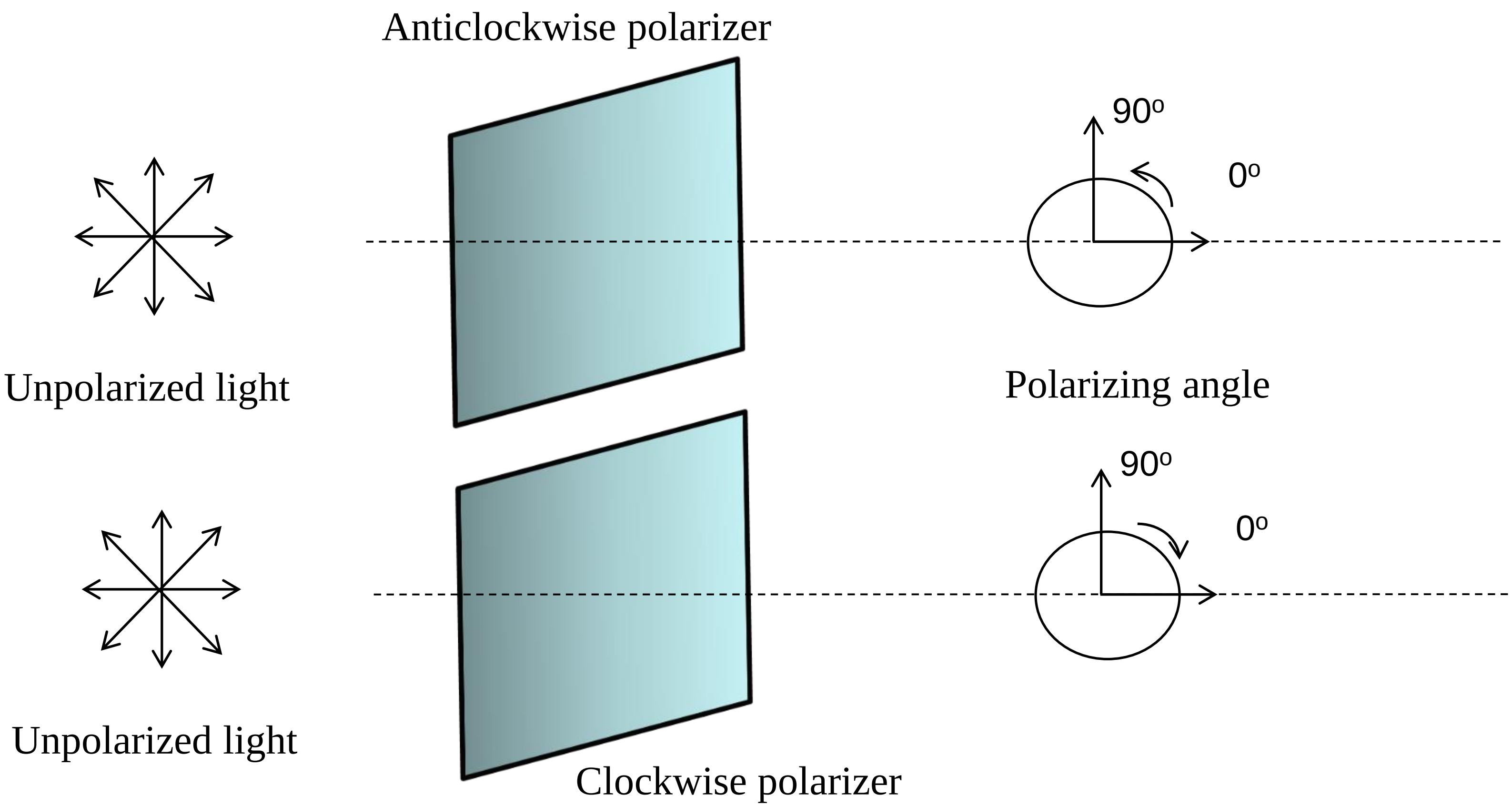
Malus's law: Intensity is attenuated by the square of the cosine of the angle between the polarization directions of the light and the polarizing filter, as

$I = I_o \cos^2 \theta.$ Hence, the viewing level must be strictly horizontal to avoid ghosting.



Circular Polarizing filter

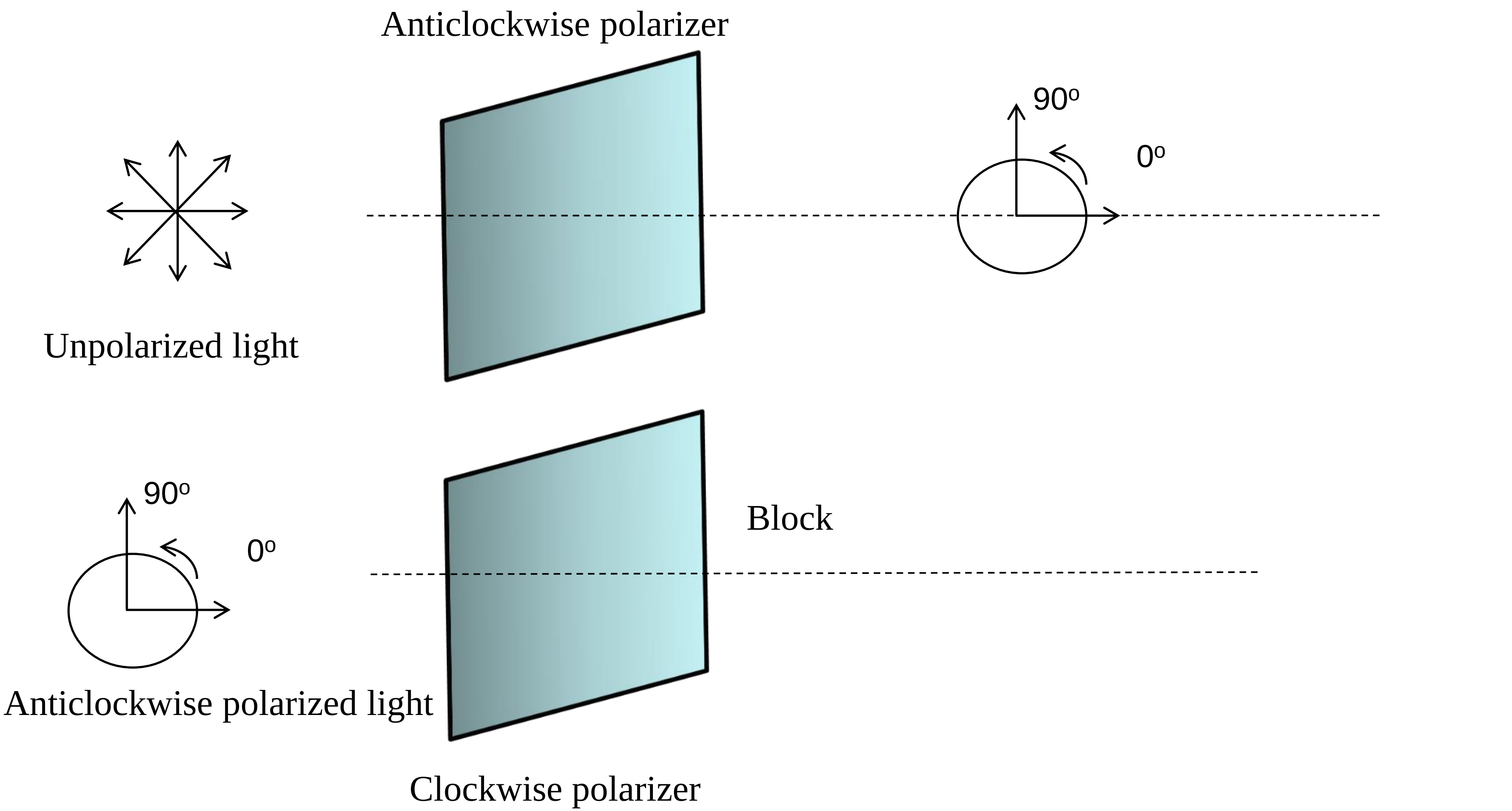
Circular Polarizing filter: polarization rotates clockwise or anticlockwise after passing through it.



http://upload.wikimedia.org/wikipedia/commons/d/d1/Circular.Polarization.Circularly.Polarized.Light_Left.Hand.Animation.305x190.255Colors.gif

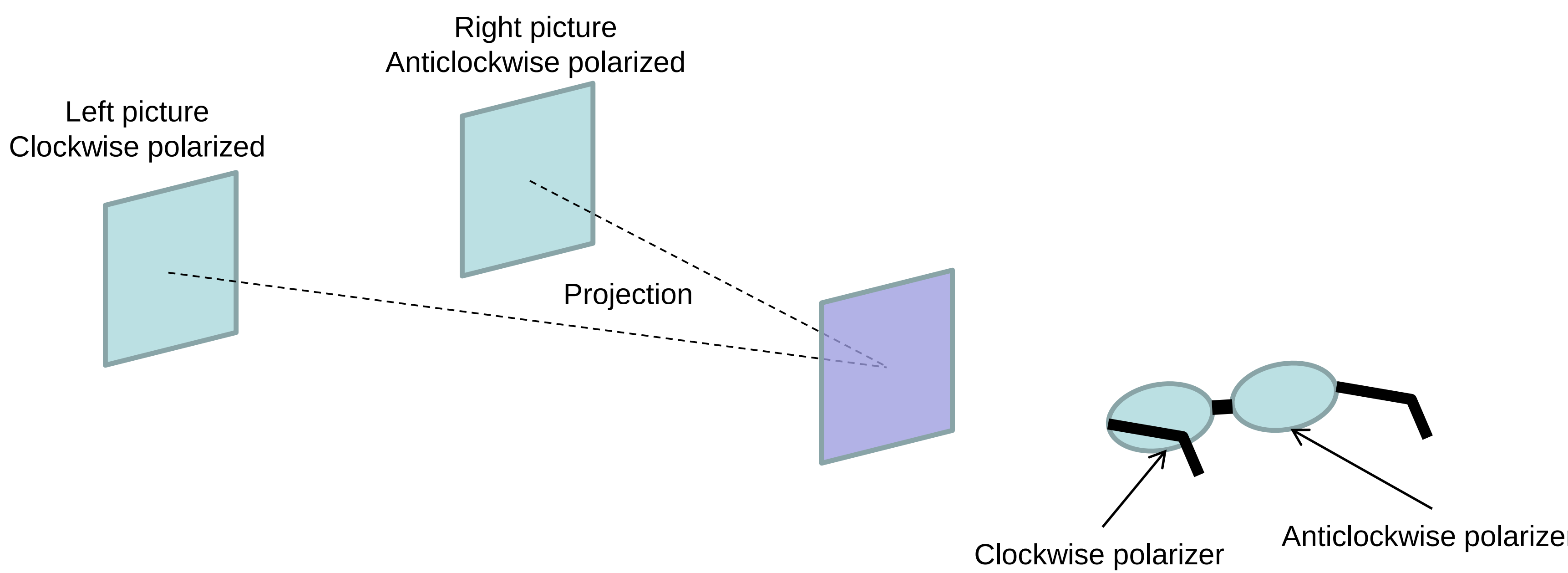
Circular Polarizing filter

Circular Polarizing filter: polarization rotates clockwise or anticlockwise after passing through it.



Circular Polarized glasses based Stereoscopic 3D

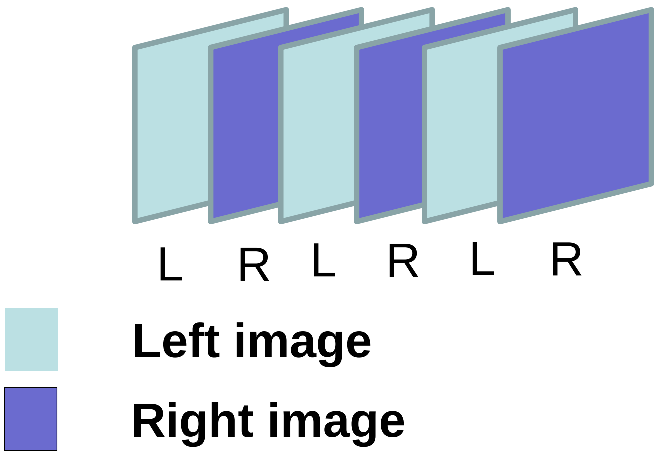
Using circular polarized images and glasses avoid the constraint on level viewing. Adopted in modern 3D movies and TV signals.

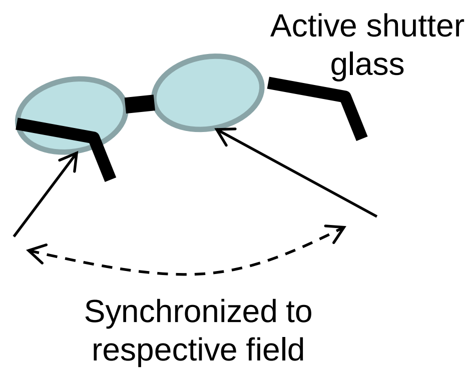


Stereoscopic 3D on TV

Time division multiplexing

Left and right images are shown in the odd and even (or even and odd) frames, respectively.



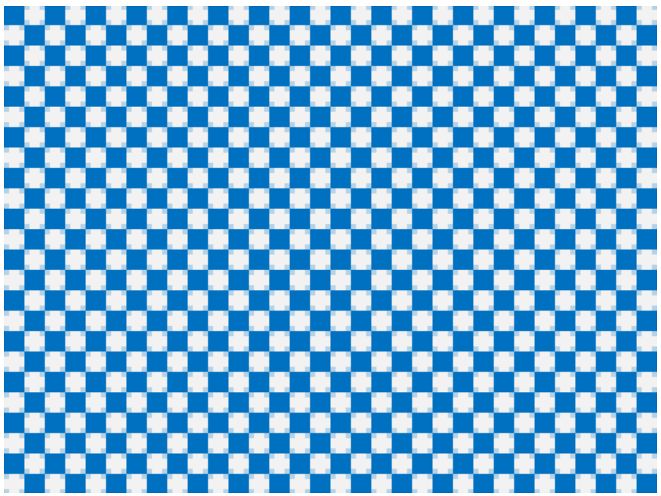


Lower brightness as glasses are only conducted at 50% of time. Expensive and bulky active shutter glasses.

Stereoscopic 3D on TV

Film-type Patterned Retarder: LG

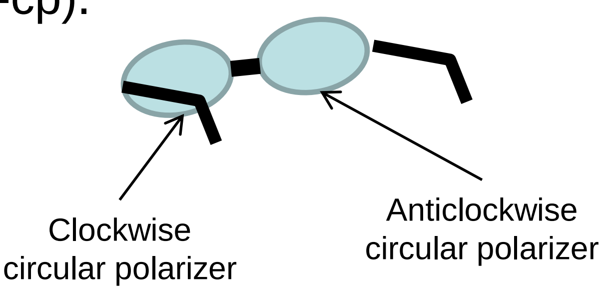
Checker sampling technology



Left eye picture (cw-cp)

Right eye picture (acw-cp)

A film with circular polarizing element is placed on each pixel. Left and right pixels have different direction of polarization (cw-cp and acw-cp).



No flickering, brighter image. Low cost glasses.

Picture resolution is lowered by a factor of 2.

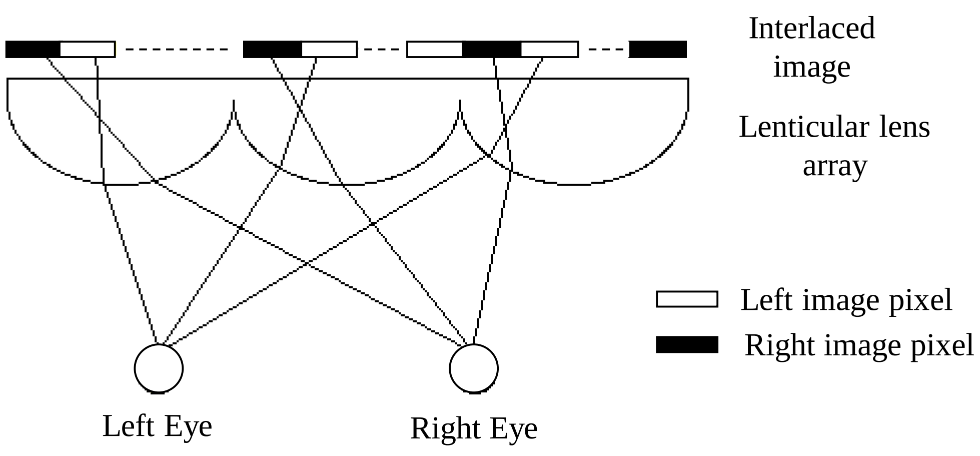
Glass free (naked eye) 3D viewing

Origin: Classic lenticular 3D photographs.

Basic concept: Interleave 2 or more views on the screen.

Overlay a lenticular lens array, or a parallax barrier.

The lens array or barrier deflects each view to its respective direction.

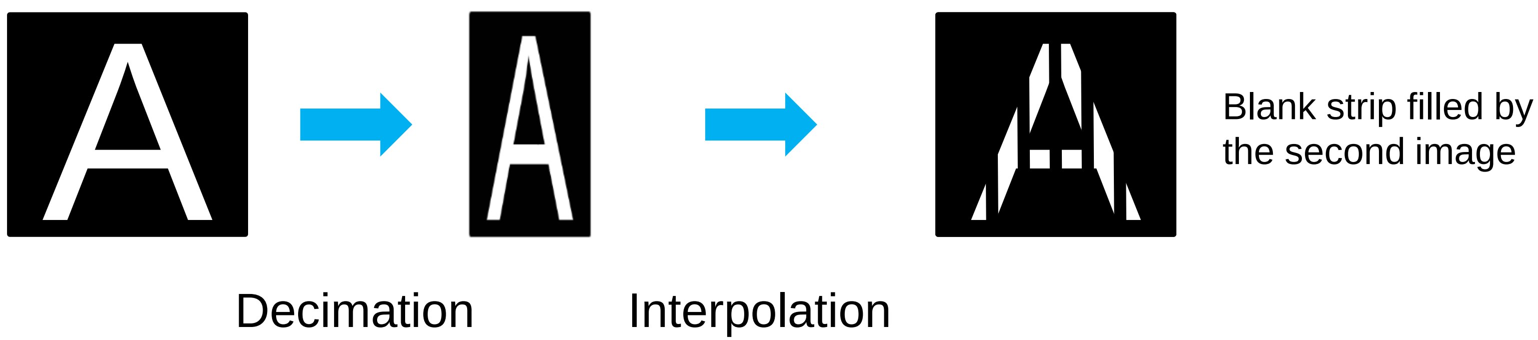


Interleaving and viewing of a pair of stereoscopic images

Glass free (naked eye) 3D viewing

Interleaving left and right images on the screen can be described as a decimation and interpolation process.

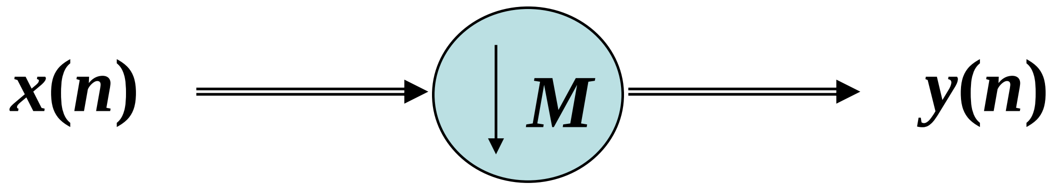
1. Each image is decimated into a new image that is half of its original width.
2. The decimated images are interpolated and interleave on the screen.



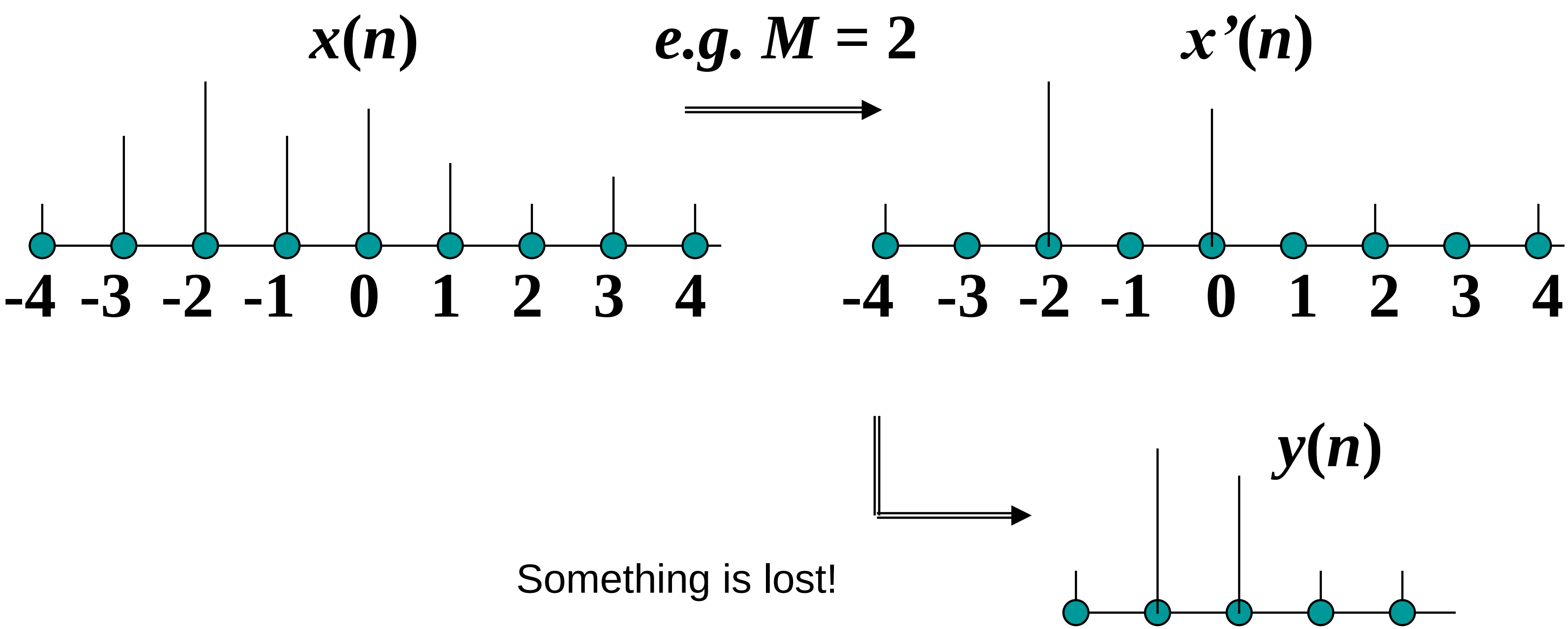
Decimation leads to distortion which cannot be fully recovered with interpolation.

Decimation, interpolation, and Interleaving of a pair of stereoscopic images

Decimation



M is the decimation factor



Decimation in 2 steps: 1st step

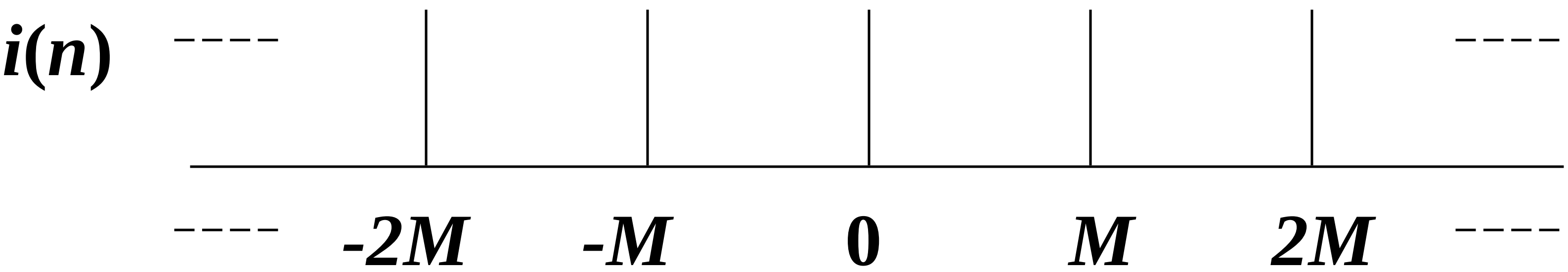
$$x(n) \Longrightarrow x'(n)$$

$$x'(n) = \begin{cases} x(n) & n = 0,\ \pm M,\ \pm 2M,\ \pm 3M,\ \pm 4M, \ldots \\ 0 & \text{otherwise} \end{cases}$$

$$x'(n) = i(n)x(n) = \left[\sum_{r=-\infty}^{\infty} \delta(n - rM)\right] x(n) \qquad (1)$$

Decimation: 1st step

$$x'(n) = i(n)x(n) = \left[\sum_{r=-\infty}^{\infty} \delta(n - rM)\right] x(n) \qquad (1)$$



***i*(*n*) is a periodic impulse train that can be expressed as**

$$i(n) = \sum_{r=-\infty}^{\infty} \delta(n - rM) = \frac{1}{M} \sum_{k=0}^{M-1} e^{j\frac{2\pi}{M}nk} \qquad (2)$$

Decimation: 1st step

$$x'(n) = i(n)x(n) = \left[\sum_{r=-\infty}^{\infty} \delta(n - rM)\right] x(n) \qquad \textbf{(1)}$$

$$i(n) = \sum_{r=-\infty}^{\infty} \delta(n - rM) = \frac{1}{M} \sum_{k=0}^{M-1} e^{j\frac{2\pi}{M}nk} \qquad \textbf{(2)}$$

$$x'(n) = \frac{x(n)}{M} \sum_{k=0}^{M-1} e^{j\frac{2\pi}{M}nk} = \frac{1}{M} \sum_{k=0}^{M-1} x(n) e^{j\frac{2\pi}{M}nk} \qquad \textbf{(3)}$$

Decimation: 2nd step

$$\boldsymbol{x(n) \Longrightarrow x'(n) \Longrightarrow y(n)}$$

$$\boldsymbol{y(n) = x'(Mn)} \qquad \textbf{(4)}$$

According to equation (3)

$$x'(n) = \frac{x(n)}{M}\sum_{k=0}^{M-1} e^{j\frac{2\pi}{M}nk} = \frac{1}{M}\sum_{k=0}^{M-1} x(n)e^{j\frac{2\pi}{M}nk}$$

$$y(n) = x'(Mn) = \frac{1}{M}\sum_{k=0}^{M-1} x(nM)e^{j2\pi nk} \qquad \textbf{(5)}$$

Decimation: 2^{nd} step

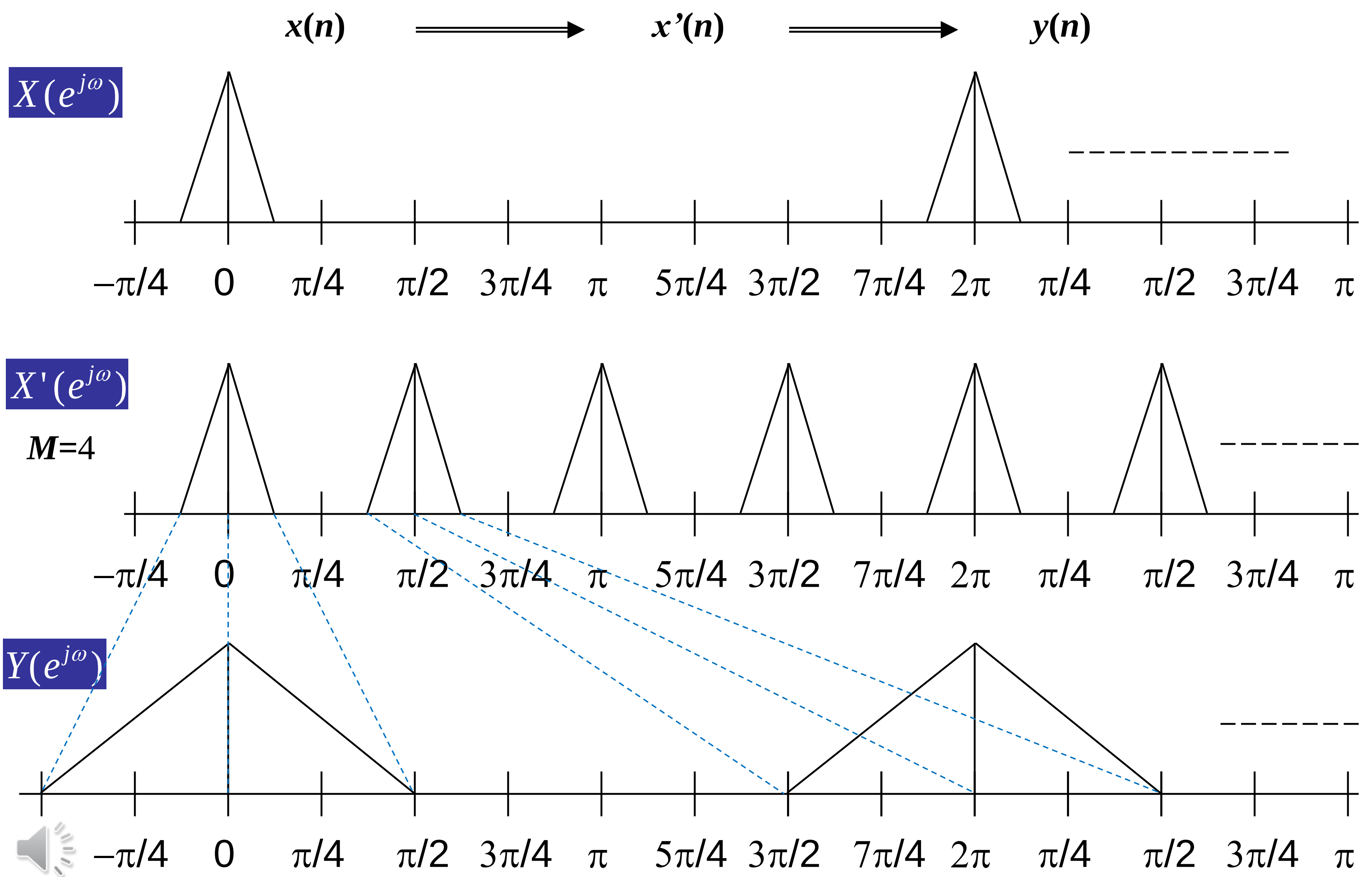
The frequency spectrum of $x(n)$ is $X\left(e^{j\omega}\right)$.

Converting the decimated signal $y(n)$ to frequency space with discrete Fourier transform:

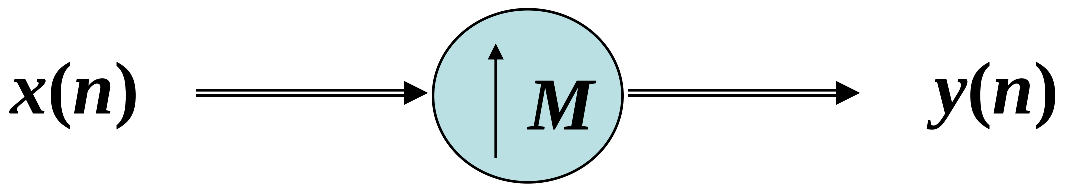
$$Y(e^{j\omega}) = \frac{1}{M}\sum_{k=0}^{M-1}\left[X\left(e^{j\left(\frac{\omega-2\pi k}{M}\right)}\right)\right]$$

1. The frequency spectrum is scaled by a factor M. Frequency axis is lengthened.
2. There are M copies (replicas) of the original base-band (scale in width).
3. The replicas are located at $\frac{2\pi k}{M}\big|_{0\le k<M}$.
4. If the bandwidth of the signal exceeds $\frac{2\pi}{M}$, the replicas overlap with each other, leading to aliasing error.

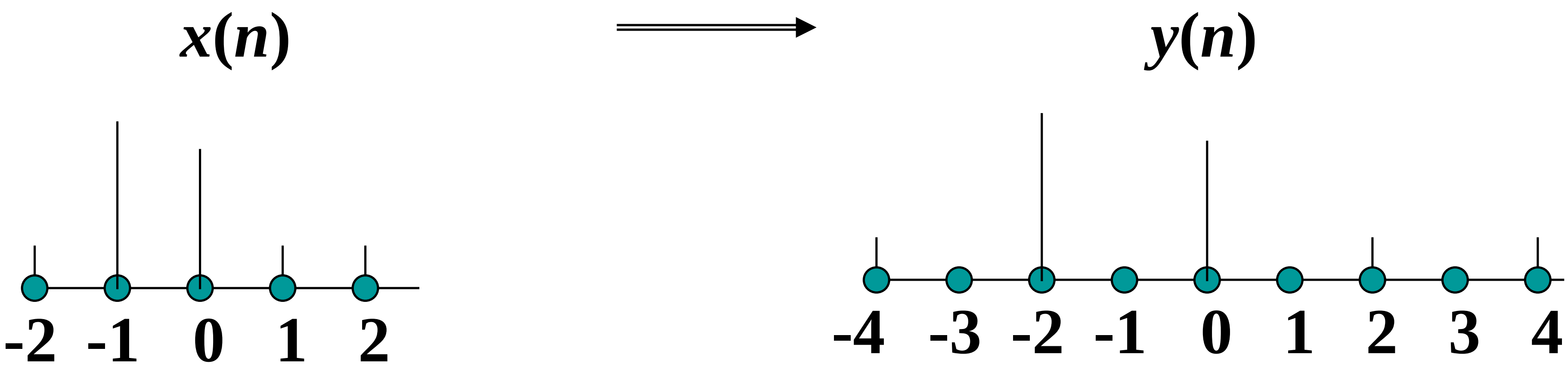
Decimation: illustration



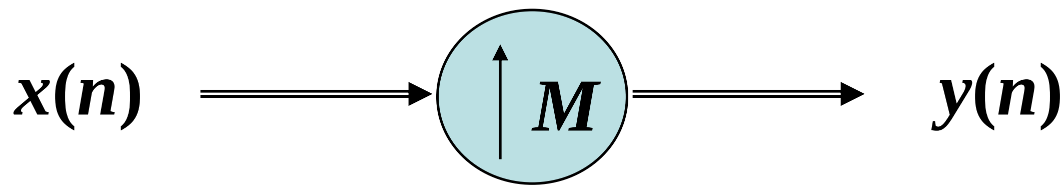
Interpolation



e.g. $M = 2$



Interpolation

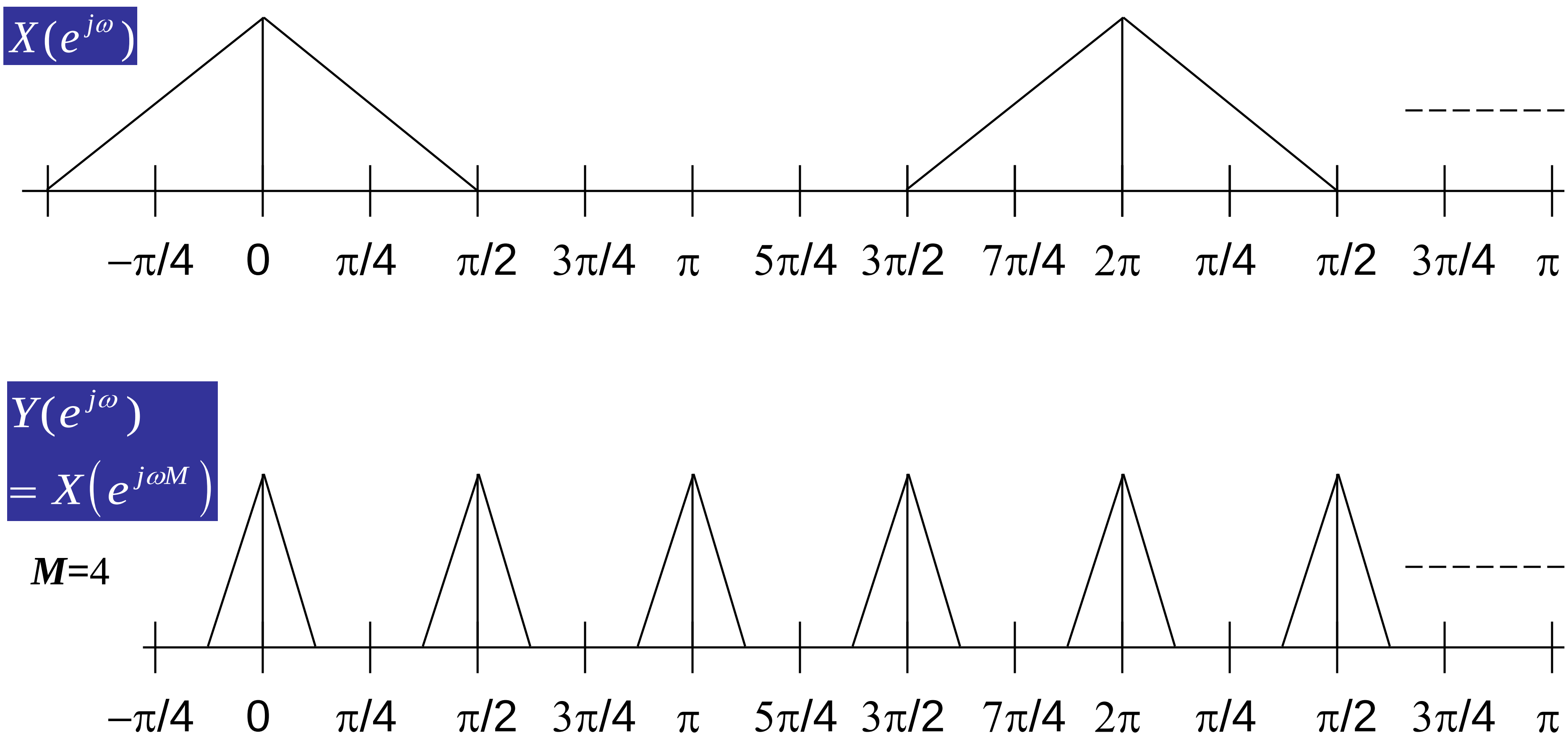


$$y(n) = \begin{cases} x(n/M) & n = 0, \pm M, \pm 2M, \pm 3M, \ldots \\ 0 & \text{otherwise} \end{cases} \qquad (6)$$

$$Y\left(e^{j\omega}\right) = X\left(e^{j\omega M}\right)$$

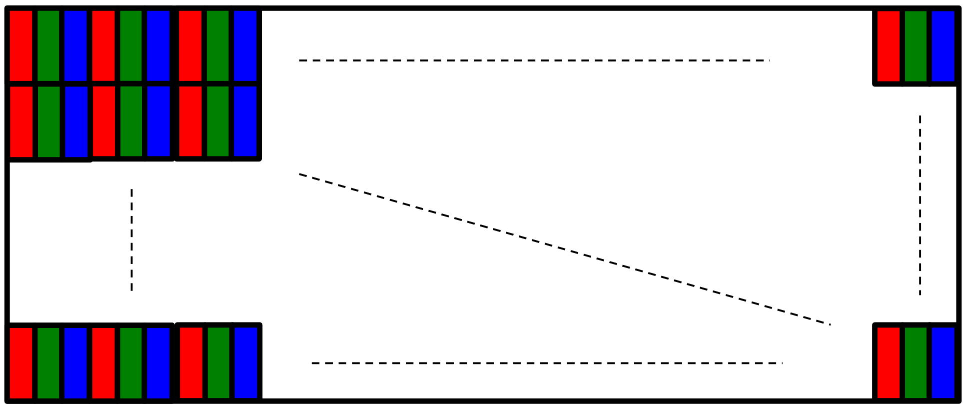
The frequency spectrum is scaled by a factor *M*. The frequency scale is compressed.

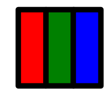
Interpolation



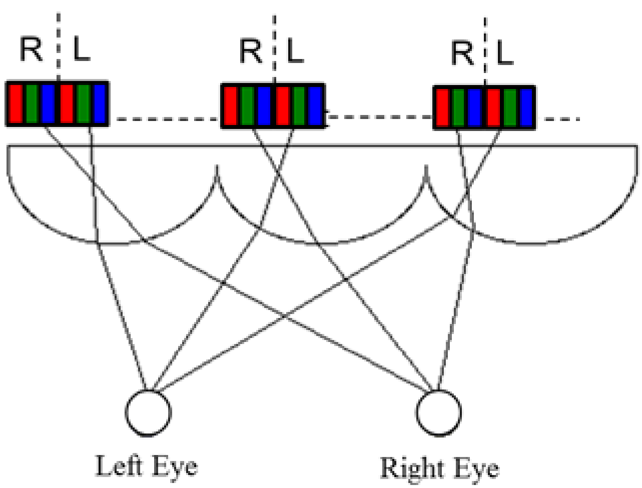
No change other than a scaling of the frequency axis.

Glass free (naked eye) 3D viewing
Does it works on monitor?





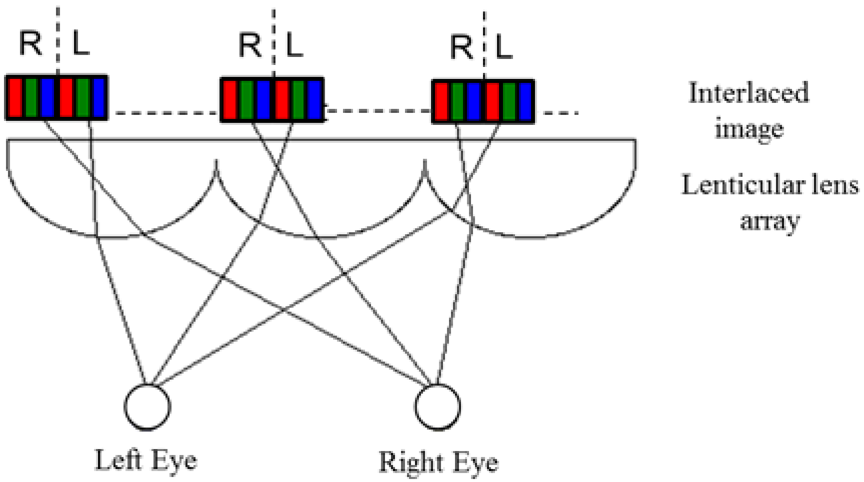
A pixel on the screen

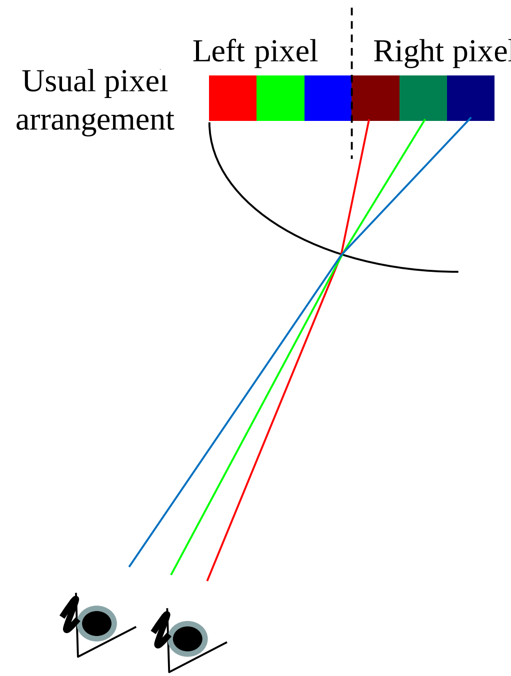


Intuitively, we can simply interleave the pixels on the monitor in alternating left and right manner.

Glass free (naked eye) 3D viewing

Early product by Sharp: Stereoscopic (2 views) with rainbow *moiré* pattern



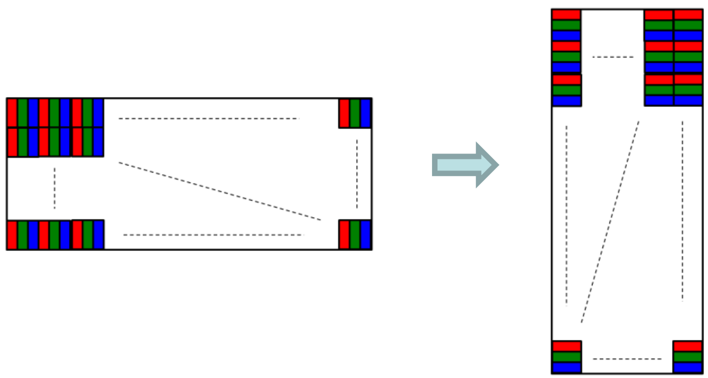


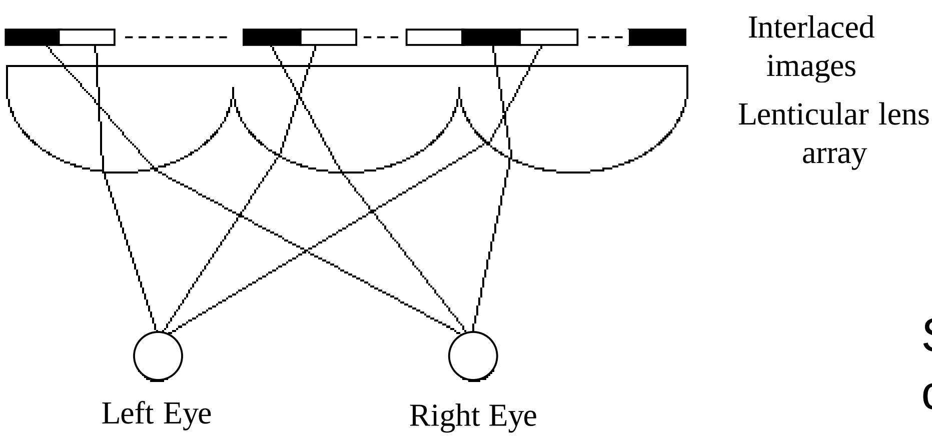
<u>Interleaving and viewing multi-view images</u>

Slight difference in viewing positions capture incomplete pixel components

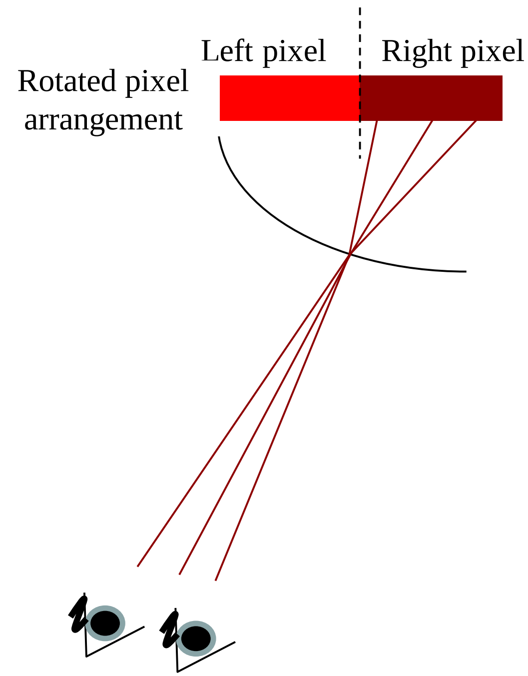
Glass free (naked eye) 3D viewing

Avoiding rainbow *moiré* patterns: Rotate screen by 90 degree.



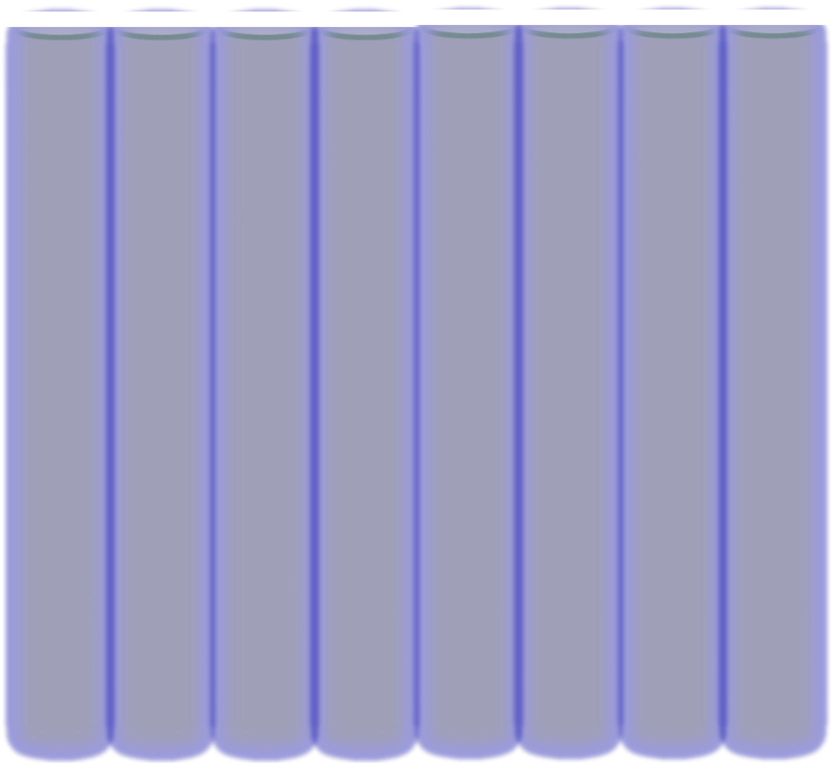


Interleaving and viewing multi-view images

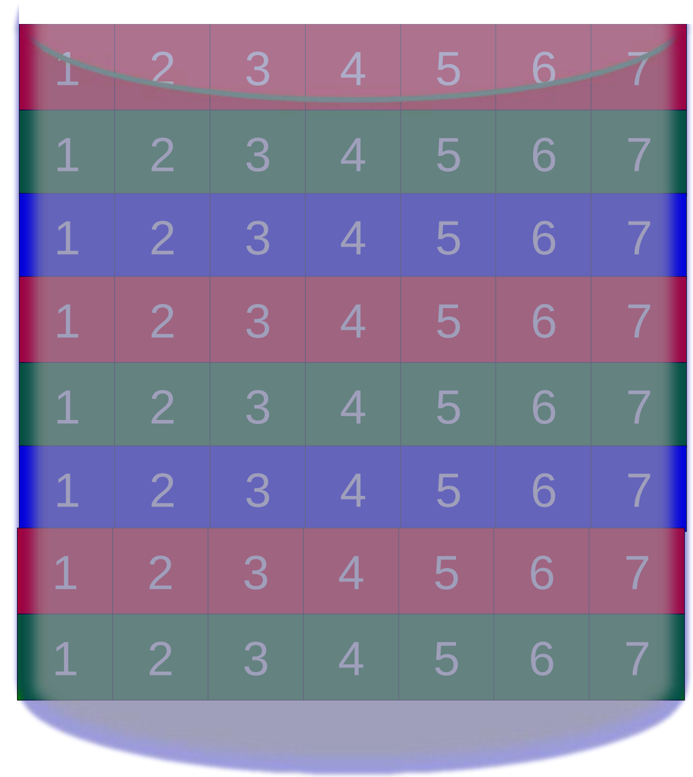


Slight difference in viewing positions capture same pixel components. Very poor resolution for large number of views.

Glass free (naked eye) 3D viewing with Auto-stereoscopic monitor



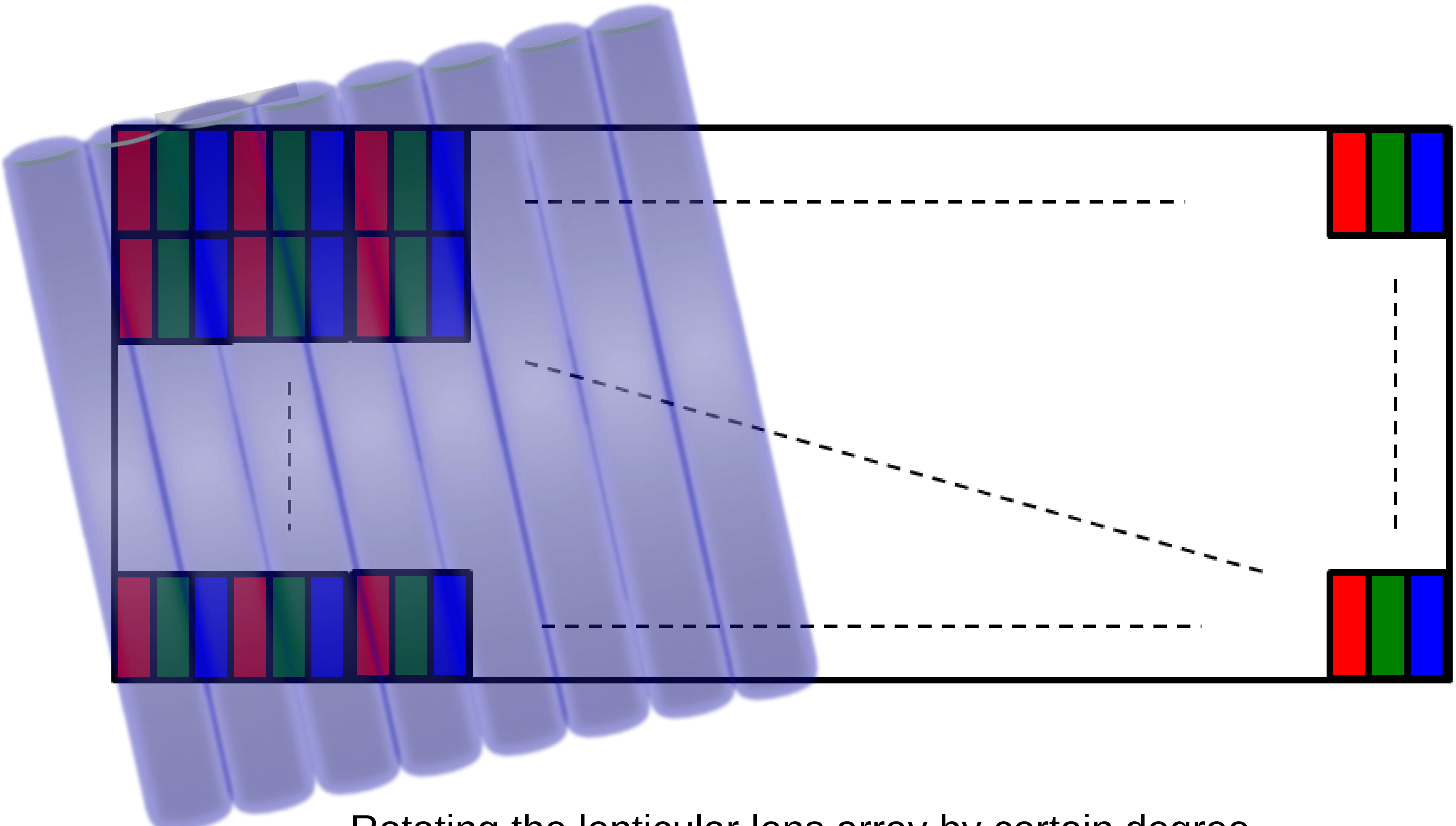
Lenticular lens array



Magnified view of one lens column

Very poor horizontal resolution

Glass free (naked eye) 3D viewing with Auto-stereoscopic monitor

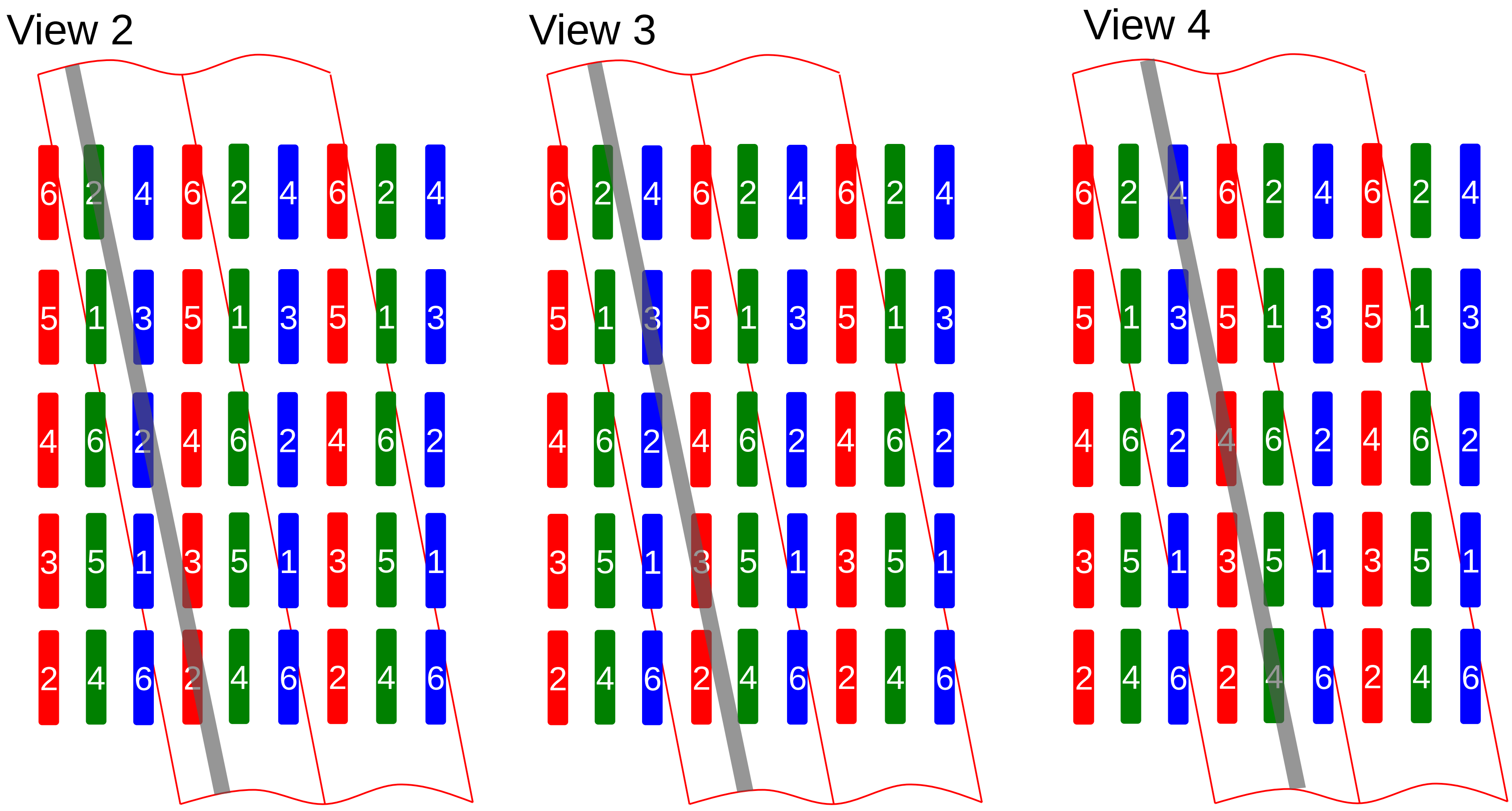


Rotating the lenticular lens array by certain degree

Glass free 3D viewing
Slant Lenticular technology

Philips patent: US6,064,424 (May 2000)

Distribute resolution degradation to horizontal and vertical directions



Packaging and distribution of muti-view signals

An early attempt to compress 3D stereoscopic sequences has been made by the MPEG with the incorporation of the Multiview Profile (MVP) [1][2] into the existing MPEG II standard. The approach is largely an extension of the temporal scalability, but so far the MVP has not been adopted as a commercially available service.

[1] M. Tanimoto and T. Fuji, “Utilization of inter-view correlation for multiple view video coding”, ISO/IEC JTC1/SC29/WG11 M11014, July 2004.

[2] H. Wang, et al, “Using inter-view prediction for multi-view video compression”, ISO/IEC JTC1/SC29/WG11 M10512, Mar. 2004.

Packaging and distribution of muti-view signals

Another approach has been developed by the Mitsubishi Electrical Research Laboratory. Its technology, known as the “simulcast” compresses individual views of video sequences with MPEG II and dispatches them via separate channel.

At the receiver they were decoded and integrated and shown on the 3D display. The method is straightforward, compatible with existing Digital Broadcasting infrastructure.

However the system is extremely complicated, requiring multiple PCs and high bandwidth networks on both the transmission and receiving sides.

Packaging and distribution of muti-view signals

The "**A**dvanced **T**hree-dimensional **Te**levision **S**ystem **T**echnologies (ATTEST) project [3] of the European Information Society Technologies (IST) aims at the development of a complete 3D-video chain including content creation, coding, transmission and display.

"Depth Image Based Rendering" (DIBR) technology is adopted where a Three Dimensional Scene is encoded as a single view plus a depth map.

To allow backward compatibility to existing 2D digital television, the base layer is encoded with MPEG II and DVB standards. The depth map is transmitted in the enhancement layer with MPEG 2/4/7 standard providing a depth value for each pixel in the base layer. At the receiver, 3D views are generated from the 2D image of the base layer and the depth map of the enhancement layer.

[3] M.O. Beeck, "Towards an Optimized 3D Broadcast Chain", *Proc. SPIE* 4864, Three-Dimensional TV, Video, and Display, 42 (November 1, 2002);

Packaging and distribution of muti-view signals

it is unlikely that the compressed multi-view 3D data could be distributed directly in the same way as the existing channels and standards.

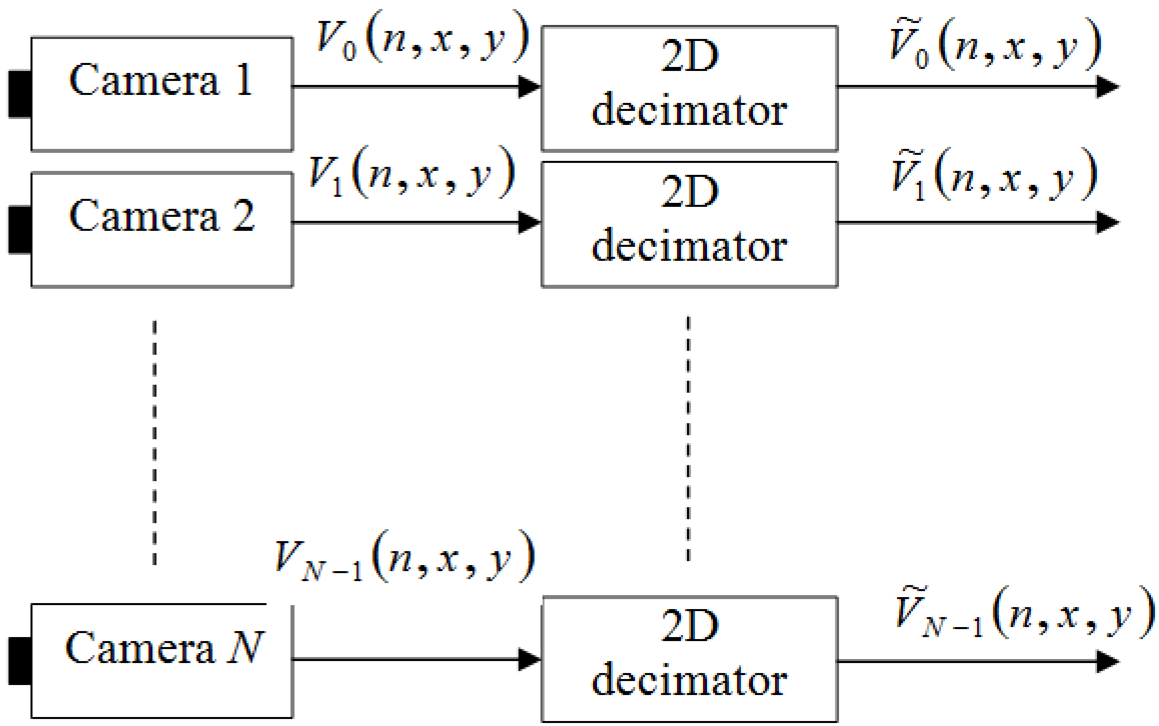
While it may be argued that technically, present standards and hardware solutions could always be changed to accommodate the newer technologies, such revolutions usually have to experience long period of incubation period before every components in the video chain could function collaboratively and survive in the consumer’s world.

StereoGraphics offers an alternative approach know as the “*N*-tiles”, a.k.a. the Hollywood Squares” format.

Images from cameras each capturing a view of the scene are down-sampled and juxtaposed as non-overlapped tiles into a single picture.

This enables multi-view images cav be processed in the same way as ordinary two dimensional pictures, and embedded into existing video chains.

Packaging and distribution of muti-view signals: N-tile format



Capturing and downsizing a video sequences with 2D decimation

Packaging and distribution of muti-view signals: N-tile format

$\tilde{V}_0(n,x,y)$	$\tilde{V}_1(n,x,y)$	$\tilde{V}_2(n,x,y)$
$\tilde{V}_3(n,x,y)$	$\tilde{V}_4(n,x,y)$	$\tilde{V}_5(n,x,y)$
$\tilde{V}_6(n,x,y)$	$\tilde{V}_7(n,x,y)$	$\tilde{V}_8(n,x,y)$

Packaging of multiple video sources into a single video sequence

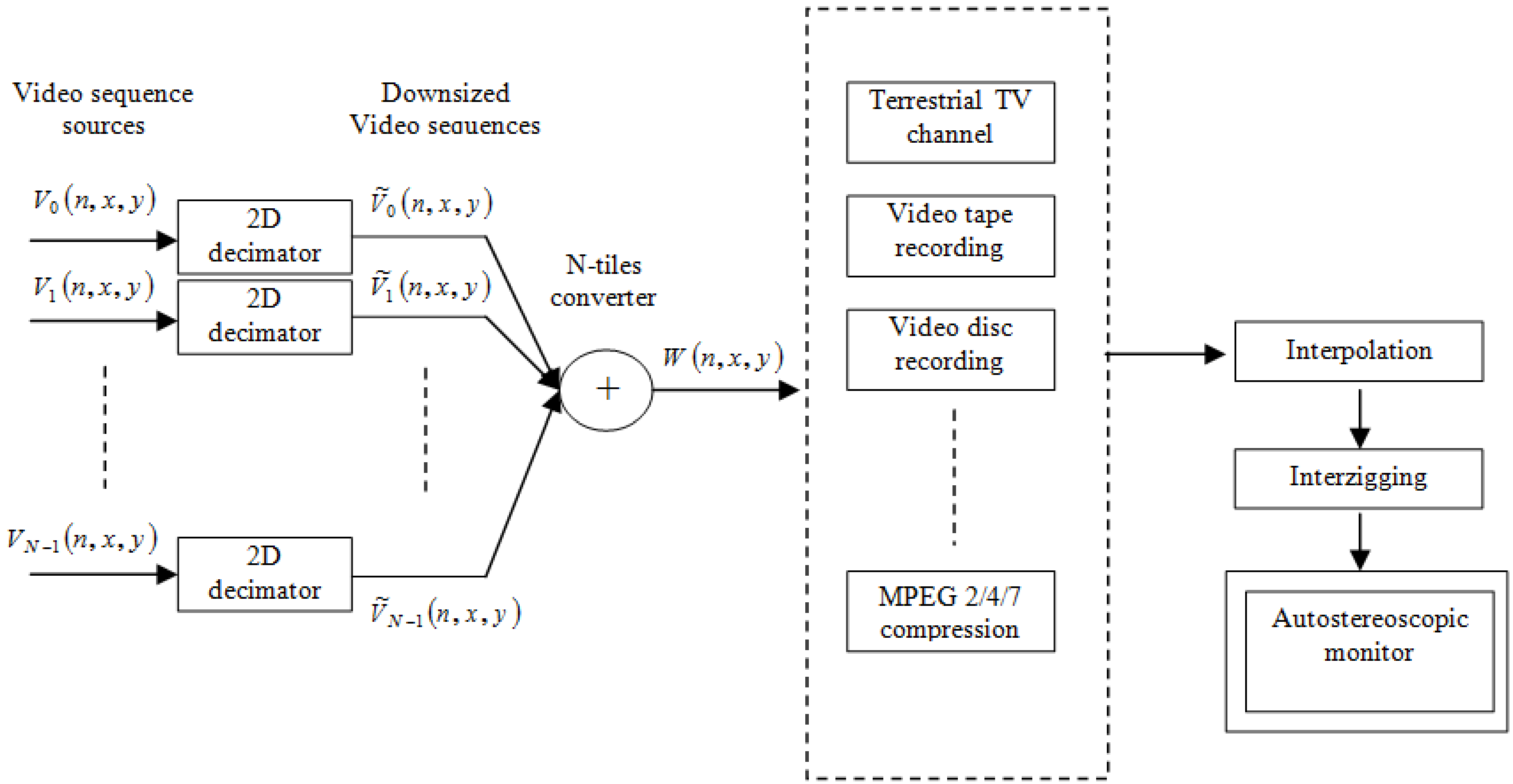
Packaging and distribution of muti-view signals: N-tile format

In general, though not mandatory, the number of cameras N is taken to be the square of an integer D, *i.e.* $D = \sqrt{N}$

$W(n,x,y)$ is just an ordinary picture frame that can be modulated, compressed, broadcast and recorded by any existing standards.

At the receiving end, each view in $W(n,x,y)$ is mapped to the corresponding sub-pixel in the auto-stereoscopic 3D monitor.

Packaging and distribution of muti-view signals: N-tile format



Integration of multi-view image frames with the *N*-tiles format