Franco Visintin

CONSIDERATIONS ON SPATIAL RESOLUTION IN 3D-TV “FRAME COMPATIBLE” FORMATS

INTRODUCTION

The need to transport 3D signals using existing HD production and distribution infrastructures has led to the development of “Dual image stereoscopic 3D imaging systems.” These systems deliver two images (left eye and right eye) that are arranged to be viewed simultaneously (or near simultaneously) by the left and right eyes. This mimics the natural binocular viewing experience and viewers perceive increased depth in the picture.

The “3D frame compatible packing method” is one of the various methods adopted for carrying the two images of a stereo pair over existing HDTV infrastructure. This method produces a 3D effect by combining the left and right pictures (a.k.a. “stereo pair”) comprising the 3D information into a single HD frame (*spatial multiplex*). Achieving this result requires downsizing the stereo pair, squeezing[1] the pictures horizontally and/or vertically.

Similar dual image stereoscopic systems were considered by the film industry. During the early "Golden Era" of 3-D cinematography of the 1950s and the second Hollywood-driven boom of 3-D films in 1983, stereoscopic motion pictures were produced through a variety of methods. Fixing a 3D lens adapter (a.k.a. *beamsplitter*) to the lens of the film camera (35mm or 16mm) enabled the camera to direct light from two separate points (left eye and right eye points), through mirrors, and on to two separate parts of the same frame of film. Two formats were used by cinema at different times:

- “Side-by-Side” storing the left and right eye images, one beside the other in a vertically split frame, squeezing them horizontally.

- “Over-Under” storing the left and right eye images, one above the other in a horizontally split frame, squeezing them vertically.

Side-by-Side on 70mm film

Over-Under on 35mm film

Today similar techniques have been adopted in the push toward “frame compatible 3D TV”:

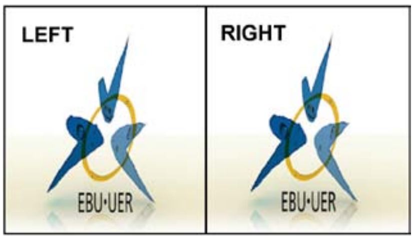
- “Side-by-Side” (SbS) = the left and right images are inserted one next to the other in the 720p or 1080i HD frame. This requires a horizontal squeeze of the video signal which halves horizontal definition.

- “Top-and-Bottom” (TaB) = the left and right images are inserted one above the other in the 720p or 1080i HD frame. Here, vertical squeeze is required which halves vertical definition.

---

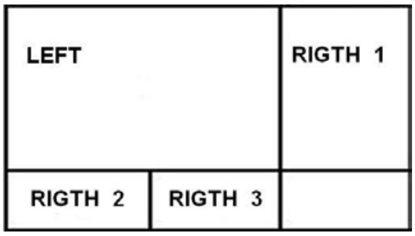
[1] *In a digital signal processing the image squeezing is performed through decimation process, proceeded by proper filtering aimed at reducing possible aliasing. The* ***decimation*** *is a technique for reducing the number of samples in a digital signal. The term comes from the Latin* decimatio *meaning "removal of a tenth" (decem = ten), a form of military discipline used by officers in the Roman Army to punish mutinous or cowardly soldiers.*

- "3D-Tile Format" = two 720p images (left and right) are packed into a single 1080p frame, inserting one 720p image unchanged, that is without any downsizing, while splitting the other 720p image into three parts (tiles) and remapping the pixel location without any filtering/decimation.



Side-by-side

Top-and-Bottom



3D Tile Format

The present paper theoretically evaluates the resolution losses of these 3D HD formats with respect to the horizontal and vertical resolution of the 2D HD formats in use in Europe today and into whose frame the stereo pair is packed: 720p/50, 1080i/25, and 1080p/50.

The same argument also applies to formats adopting 59.94/60 rates.

RESOLUTION EVALUATION

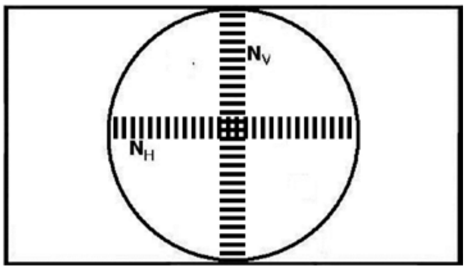
In cinema and television the term "resolution" is understood to evaluate how close picture details (ideally considered as *alternate bright and dark lines*) can be to each other and still be visibly resolved.

In cinematography, resolution evaluation is tied to physical sizes, specified as the "number of lines per unit-distance" (L/mm, lines per mm) along the vertical and horizontal axis.

In television, resolution evaluation is related to the overall size of a picture, specified horizontally and vertically by the "number of lines per picture height" (LPH). In particular:

- The vertical resolution N_V defines the capability of the system to resolve horizontal lines and depends primarily on the number of scanning lines per picture and the combined effects of the camera and display capabilities (Kell and Interlace effects, described in greater detail below). It is specified as the number of distinct horizontal lines which can be satisfactorily resolved on the television screen.

- The horizontal resolution N_H defines the capability of the system to resolve vertical lines. It depends on the sampling frequency adopted by the system,, taking into consideration the aspect ratio (AR) of the system and the cut-off frequency of the anti-aliasing filter. It is expressed as the number of distinct vertical lines, which can be satisfactorily resolved in 1/AR of the width of a television screen.[2]



[2] *To maintain the same spatial scale for vertical and horizontal resolution, horizontal resolution is also specified as "lines per picture height" (LPH) rather than "lines per picture width" (LPW).*

VERTICAL RESOLUTION

The vertical resolution would equal, only ideally, the "number of active lines" (Nal) per frame. This would happen only if the scanning lines of any acquisition means (camera or scanner, where the scanning line is performed by a row of sensitive cells) were centered on the picture details. Conversely, complete loss of vertical resolution will occur when the scanning spot straddles the picture details.

A loss in the vertical resolution can be signified by multiplying the number of active lines (Nal) by the vertical Kell factor (vKf), whose value, which can be statistically detected with progressive scanning, was between 0.6 and 0.8 for the old cathode ray sensors and displays and between 0,85 and 0,95[3] for modern fixed pixel scanners (CCD, CMOS) and displays (LCD, PDP).

The value of the Vertical Resolution is therefore:

$$\mathbf{N_V = Nal \times vKf} \quad \text{[lines]}$$

Another vertical resolution loss could arise only on moving details in the image for interlaced scanning systems. Such loss is signified by the Interlace factor (If), described as "*the ratio of lines of resolution as perceived in a video picture produced using interlaced scan, divided by the lines of resolution as perceived in the same video picture except as produced using progressive scan*" (see Ref.1). It is statistically evaluated as 0.7, by which the active lines should be multiplied to obtain the complete interlaced vertical resolution N_{Vi}:

$$\mathbf{N_{Vi} = Nal \times vKf \times If} \quad \text{[lines]}$$

Note: for progressive pictures the Interlace factor value is 1.0.

HORIZONTAL RESOLUTION

The horizontal resolution would equal, only ideally, the "number of samples per active line" (Nsa) divided by the aspect ratio (AR). This would happen only if the cells of the sensors of the acquisition means (camera, scanner) were horizontally centred on the picture details. Conversely, complete loss of horizontal resolution will occur when the sensor cell straddles the picture details. This loss is signified by the horizontal Kell factor (hKf), whose value is between 0.85 and 0.95.

The value of the Horizontal Resolution is therefore:

$$\mathbf{N_H = Nsa \times hKf \;/\; AR} \quad \text{[lines]}$$

The following Table 1 summarizes the values of the vertical N_{V2D} and horizontal N_{H2D} resolution calculated for the 2D HD systems presently available.

[3] *From subjective data, obtained with progressive (non-interlaced) scanning, RCA engineer Raymond D. Kell found in 1934 that vertical resolution is reduced to 64% of the number of active lines. This value was later revised: to 0.7 for electron beam scanning (in analog TV) and more recently to 0.85 - 0.95, when fixed pixel scanning (e.g., CCD or CMOS) and fixed pixel displays (e.g. LCD or plasma) entered the market.*

PARAMETER	unit	720p50	1080i25	1080p50
Frame rate (frames per second)	fps	50	25	50
Number of active lines per frame Nal	---	720	1080	1080
Number samples per active line Nsa	---	1280	1920	1920
Vertical Kell factor vKf	---	0.90	0.90	0.90
Horizontal Kell factor hKf	---	0.90	0.90	0.90
Interlace factor If	---	1.00	0.70	1.00
Vertical resolution (lines per picture height) N_{V2D}	LPH	≈648	≈680	≈972
Horizontal resolution (lines per picture height) N_{H2D}	LPH	≈648	≈972	≈972

TABLE 1 - Parameters of 2D HD Systems

EVALUATION OF THE H & V RESOLUTION OF THE 3DTV PICTURES

Packing 3D information comprised of two images inside a single frame is accomplished by squeezing the images horizontally (reducing the number of samples per active line N_{sa}), vertically (reducing the number of active lines N_{al}) or both.

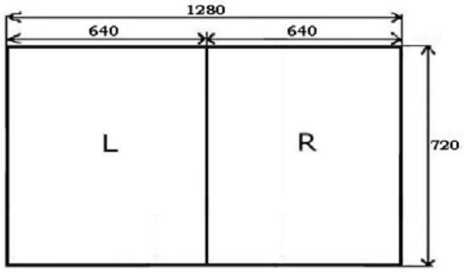
This squeeze reduces the corresponding resolution by the following factors:
- vertical reduction factor Vrf = reduced N_{al} / original N_{al}
- horizontal reduction factor Hrf = reduced N_{sa} / original N_{sa}

So that the H and V resolutions of the 3D left and right images are:
- 3D vertical resolution $\mathbf{N_{V3D}}$ = N_{V2D} x Vrf
- 3D horizontal resolution $\mathbf{N_{H3D}}$ = N_{H2D} x Hrf

Having established a baseline for 2D image resolution, the foregoing calculates the new values for the vertical resolution N_V and horizontal resolution N_H under various packing conditions that have been adopted by the 3D industry taking into account the values of N_{V2D} e N_{H2D} of Table 1.

The values of Nv and N_H described below provide interesting insight on the performance and usability of various 3D systems.

3D SIDE-by-SIDE 720p/50



The left and right images, horizontally squeezed to a 640x720 aspect ratio, are packed side-by-side inside a 1280x720 frame and transmitted in the 720p/50 format.

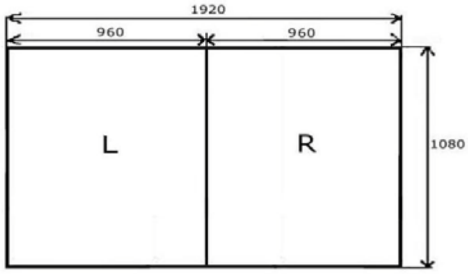
Under these conditions

Vrf = 720/720 = 1. 0 Hrf = 640/1280 = 0.5

the resolution values become:

Vertical resolution $\mathbf{N_{V3D}} = N_{V2D} \times Vrf = \approx 648 \times 1.0 = \approx 648$ LPH
Horizontal resolution $\mathbf{N_{H3D}} = N_{H2D} \times Hrf = \approx 648 \times 0.5 = \approx 324$ LPH

3D SIDE-by-SIDE 1080i/25



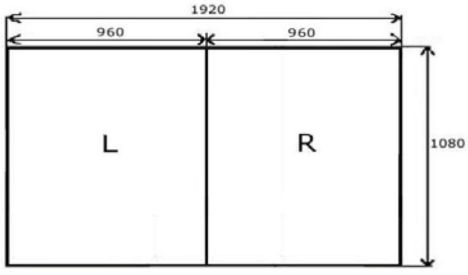
The left and right images, horizontally squeezed to a 960x1080 aspect ratio, are packed side-by-side inside a 1920x1080 frame and transmitted in the 1080i/25 format.

Under these conditions
$Vrf = 1080/1080 = 1.0$ $Hrf = 960/1920 = 0.5$

the resolution values become:

Vertical resolution $\mathbf{N_{V3D}} = N_{V2D} \times Vrf = \approx 680 \times 1.0 = \approx 680$ LPH
Horizontal resolution $\mathbf{N_{H3D}} = N_{H2D} \times Hrf = \approx 972 \times 0.5 = \approx 486$ LPH

3D SIDE-by-SIDE 1080p/50



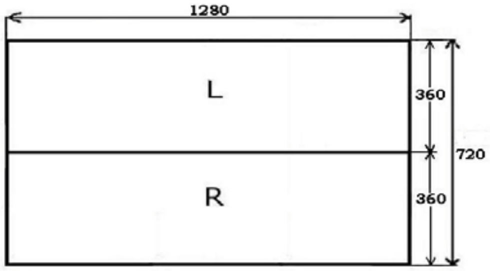
The left and right images, horizontally squeezed to a 960x1080 aspect ratio, are packed side-by-side inside a 1920x1080 frame and transmitted in the 1080p/50 format.

Under these conditions
$Vrf = 1080/1080 = 1.0$ $Hrf = 960/1920 = 0.5$

the resolution values become :

Vertical resolution $\mathbf{N_{V3D}} = N_{V2D} \times Vrf = \approx 972 \times 1.0 = \approx 972$ LPH
Horizontal resolution $\mathbf{N_{H3D}} = N_{H2D} \times Hrf = \approx 972 \times 0.5 = \approx 486$ LPH

3D TOP-and-BOTTOM 720p/50



The left and right images, vertically squeezed to a 360x1280 aspect ratio, are packed “top-and bottom” inside a 1250x720 frame and transmitted in the 720p/50 format.

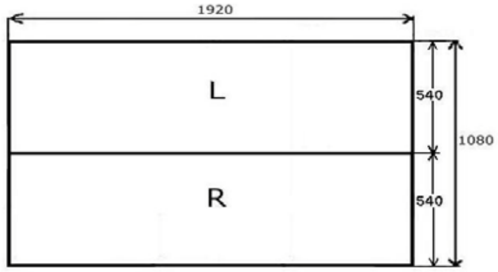
Under these conditions

Vrf = 360/720 = 0.5   Hrf = 1280/1280 = 1.0

the resolution values become:

Vertical resolution $\mathbf{N_{V3D}} = N_{V2D} \times Vrf = \approx 648 \times 0.5 = \approx 324$ LPH
Horizontal resolution $\mathbf{N_{H3D}} = N_{H2D} \times Hrf = \approx 648 \times 1.0 = \approx 648$ LPH

3D TOP-and-BOTTOM   1080i/25



The left and right images, vertically squeezed to a 1920x540 aspect ratio, are packed “top-and bottom” inside a 1920x1080 frame and transmitted in the 1080i/25 format.

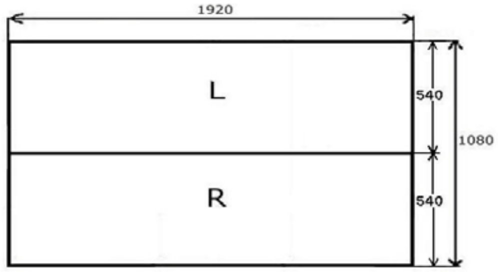
Under these conditions

Vrf = 540/1080 = 0.5   Hrf = 1920/1920 = 1.0

the resolution values become:

Vertical resolution $\mathbf{N_{V3D}} = N_{V2D} \times Vrf = \approx 680 \times 0.5 = \approx 340$ LPH
Horizontal resolution $\mathbf{N_{H3D}} = N_{H2D} \times Hrf = \approx 972 \times 1.0 = \approx 972$ LPH

3D TOP-and-BOTTOM   1080p/50



The left and right images, vertically squeezed to a 1920x540 aspect ratio, are packed “top-and bottom” inside a 1920x1080 frame and transmitted in the 1080i/25 format.

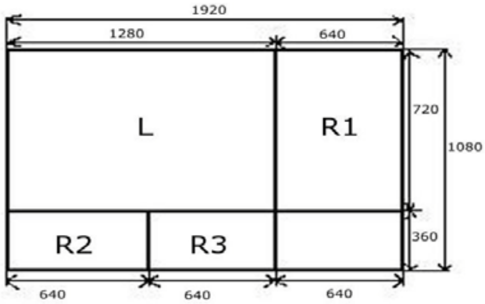
Under these conditions

Vrf = 540/1080 = 0.5   Hrf = 1920/1920 = 1.0

the resolution values become:

Vertical resolution $\mathbf{N_{V3D}} = N_{V2D} \times Vrf = \approx 972 \times 0.5 = \approx 486$ LPH
Horizontal resolution $\mathbf{N_{H3D}} = N_{H2D} \times Hrf = \approx 972 \times 1.0 = \approx 972$ LPH

3D TILE FORMAT   1080p/50



The left and right images, both in the 720p/50 format, are packed inside a 1920x1080 frame and transmitted in the 1080p/50 format. The squeezing process is replaced by a much safer re-mapping process which eliminates image degradation problems.

The left image is inserted in the upper left part of the 1920x1080 frame while the right image is split into three tiles inserted in the remaining space of the 1920x1080 frame.

Accordingly, the resolution is the same as in the case of the 720p format.

Vertical resolution $\mathbf{N_{V3D}}$ = ≈648 LPH
Horizontal resolution $\mathbf{N_{H3D}}$ = ≈648 LPH

CONCLUSIONS

The following Table 2 summarizes the values of the vertical N_V and horizontal N_H resolution calculated for the 2D and 3D HD systems presently available. The ≈ sign reminds us that the resolution values are highly dependent on the values attributed to the “Kell factors” (horizontal and vertical) and to the “Interlace factor” (a source of ongoing debate since their discovery by Raymond D. Kell in 1934).

Whether the human eye is more sensitive to horizontal resolution or to vertical resolution is debated in today’s literature. In any case, the data presented in Table 2 for the SbS and TaB formats show the significant divergence between the two resolutions. The overall assessment of the viewer might be influenced by the lowest of the two resolutions: further investigations would be needed in order to confirm this assumption. No doubt, however, that the best formats are those who preserve the balance between the two resolutions.

PARAMETER	units	2D formats			3D formats		
		720p50	1080i25	1080p50	SbS 1080i25	TaB 720p50	3D Tile 1080p50
Frame rate (frames per second)	fps	50	25	50	25	50	50
Number of active lines per frame N_{al}	--	720	1080	1080	1080	720	720
Number of samples per active line N_{sa}	--	1280	1920	1920	1920	1280	1280
Vertical resolution (samples per picture heigth) N_V	SPH	648	680	972	680	324	648
Horizontal resolution (samples per picture heigth) N_H	SPH	648	972	972	486	648	648

TABLE 2 - Vertical and Horizontal Resolution in 2D and 3D Systems

REFERENCES

1. Allan W. Jayne, Jr., "Video and Scanner Resolution -- The Kell Factor", www.cockam.com/kell.htm, 1997-2000

2. ITU-R, "Parameter values for the HDTV standards for production and international programme exchange", Rec. ITU-R BT.709-5 (04/2002)

3. Michael Robin, "Revisiting Kell", Broadcast Engineering, March 1, 2003

4. Michael Robin, "Horizontal Resolution: Pixel or Lines", Broadcast Engineering, April 1, 2005

5. Steve Mullen, Just What is 1080i? , HDV@Work, Feb. 2006 http://digitalcontentproducer.com/hdhdv/depth/hdv_at_work_02272006/

6. Hans Hoffmann, "HDTV - EBU format comparisons at IBC-2006", EBU Technical Review – October 2006

7. EBU-UER, "High Definition (HD) Image Formats for Television Production", EBU-Tech 3299, Geneva January 2010

Franco Visintin, Chairman of the Italian Section of SMPTE;
Former Chief Engineer, RAI TV Production Center of Milan (Italy)