Quantitative Evaluation of Human Visual Perception for Multiple Screens and Multiple CODECs

Sean McCarthy, Ph.D.
Fellow of the Technical Staff
Motorola Mobility

SMPTE 2012

New Opportunities

More Form Factors

More Resolutions – qHD to UHDTV

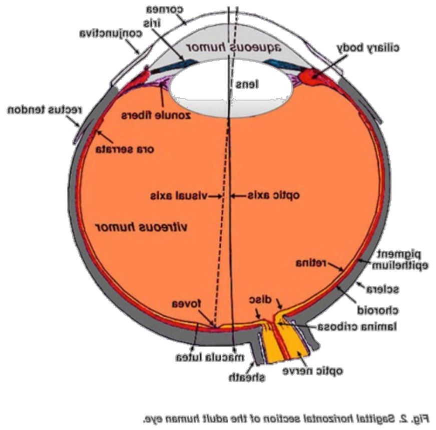
More CODECs – AVC and HEVC

How can we make sure we continue to deliver great consumer experiences?

New Challenges

How will consumers judge video quality?

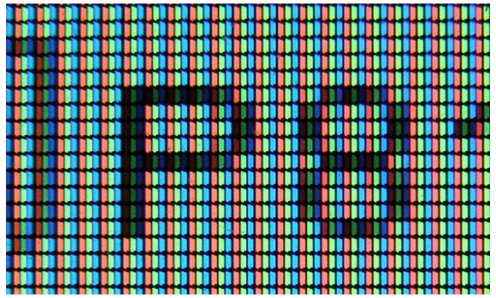
What bandwidth will we need?

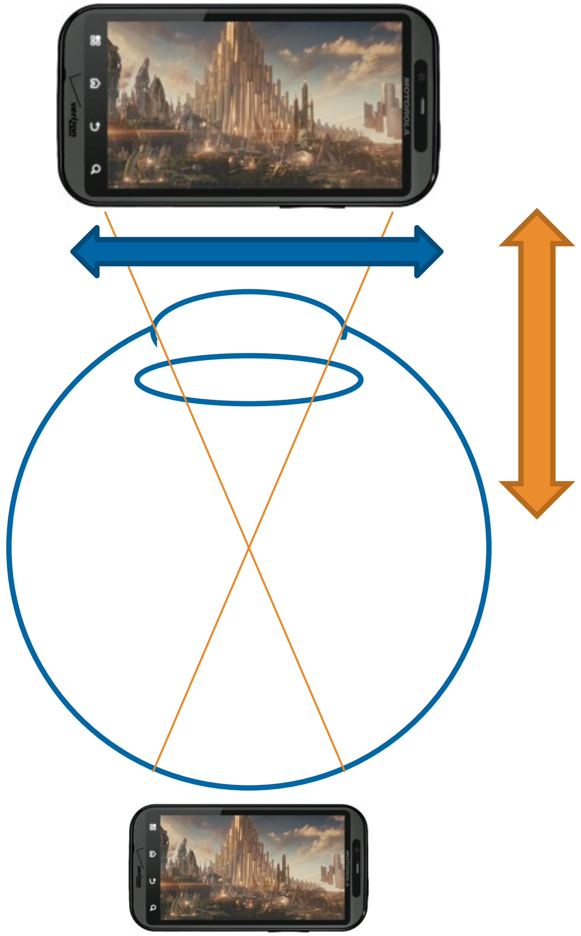


Can Applied Vision Science provide any Answers?

SMPTE 2012

Acuity & Artifacts

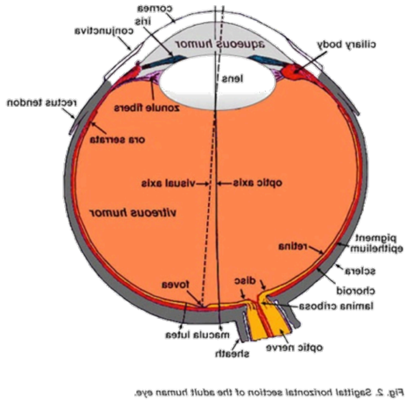


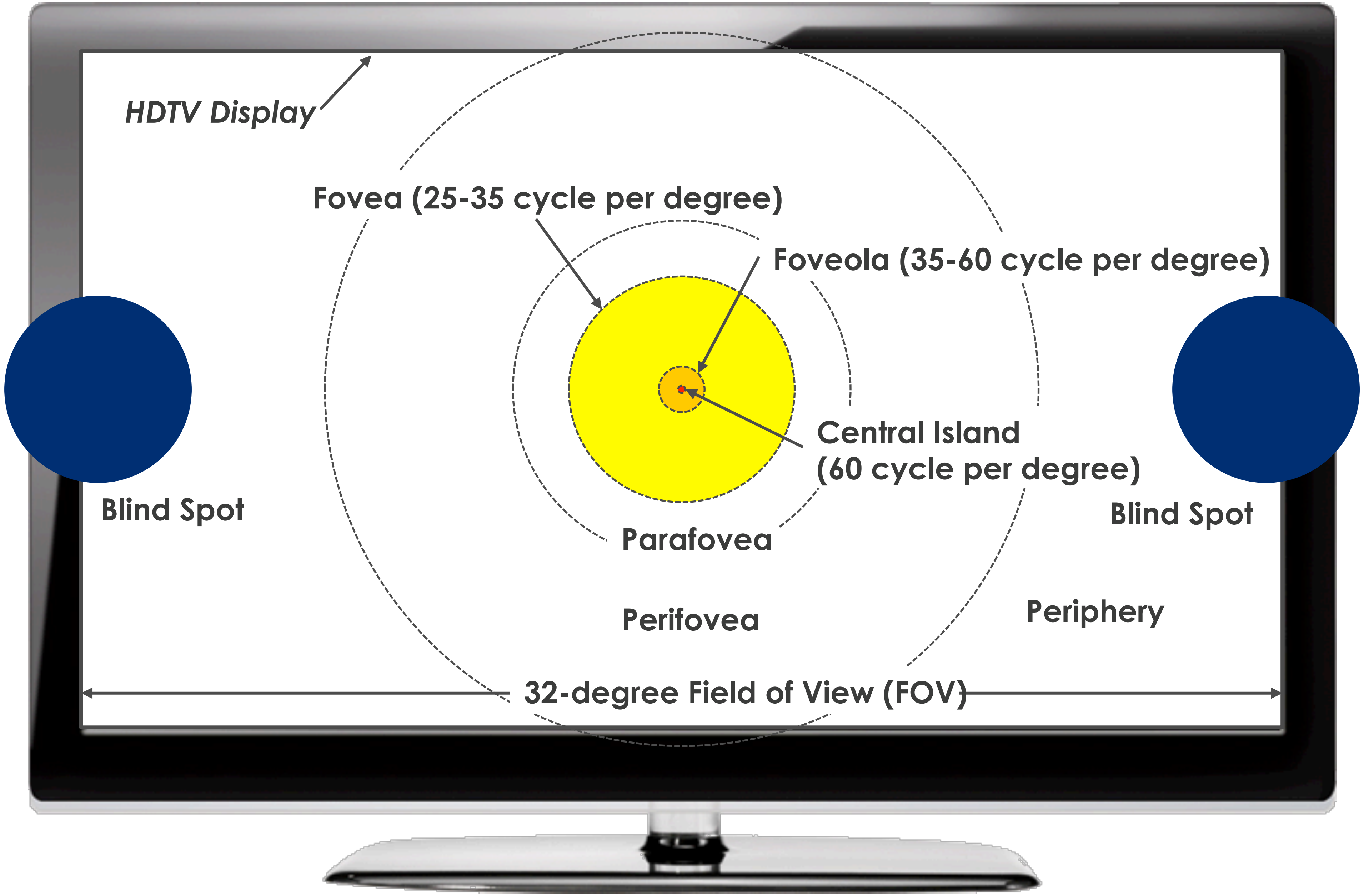


Perceived Video Quality is a function of Visual Acuity

Pixels per degree = Display Resolution / Field of View (FOV)

Topography of photoreceptor layer

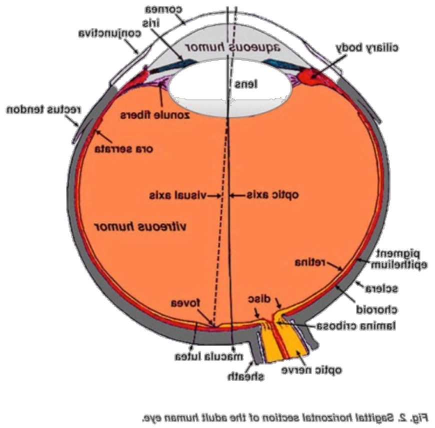


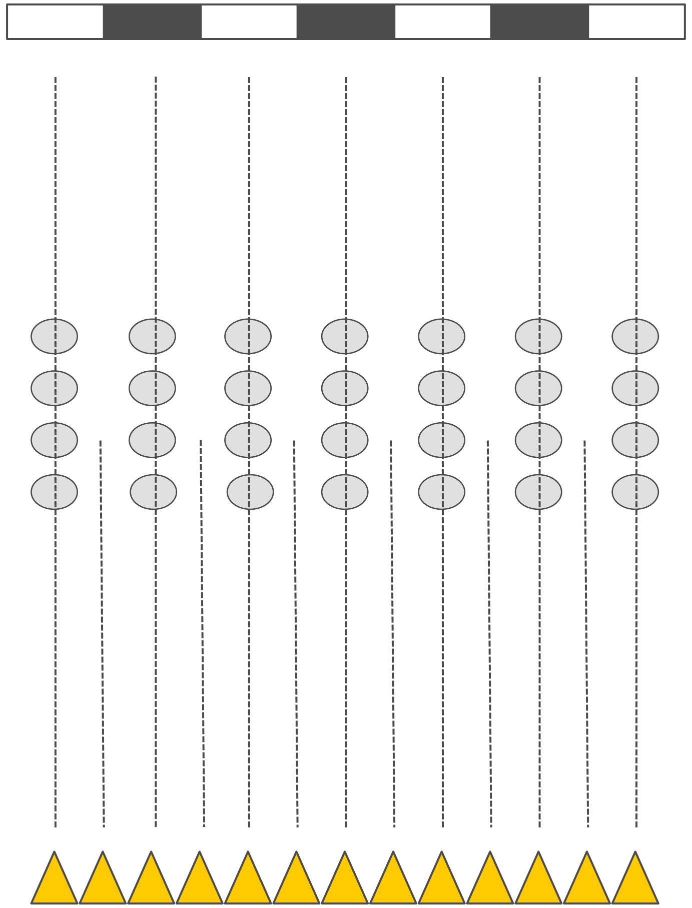


Maximum acuity = 60 cycles per degree

We build up our mind's eye image a bit at time.

Image & Display Resolution Limits



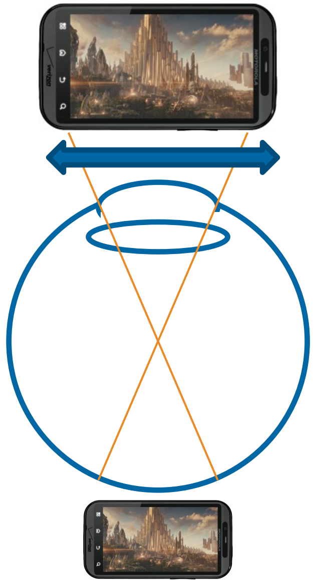


Nyquist Limit for Image on Display
30 cycles per degree

Pixel Sampling
(Nyquist Limit for Retina)
60 cycles per degree

Cone Sampling
120 cones per degree

There's more than one Type of Acuity



Hyperacuity → "Feature Localization"

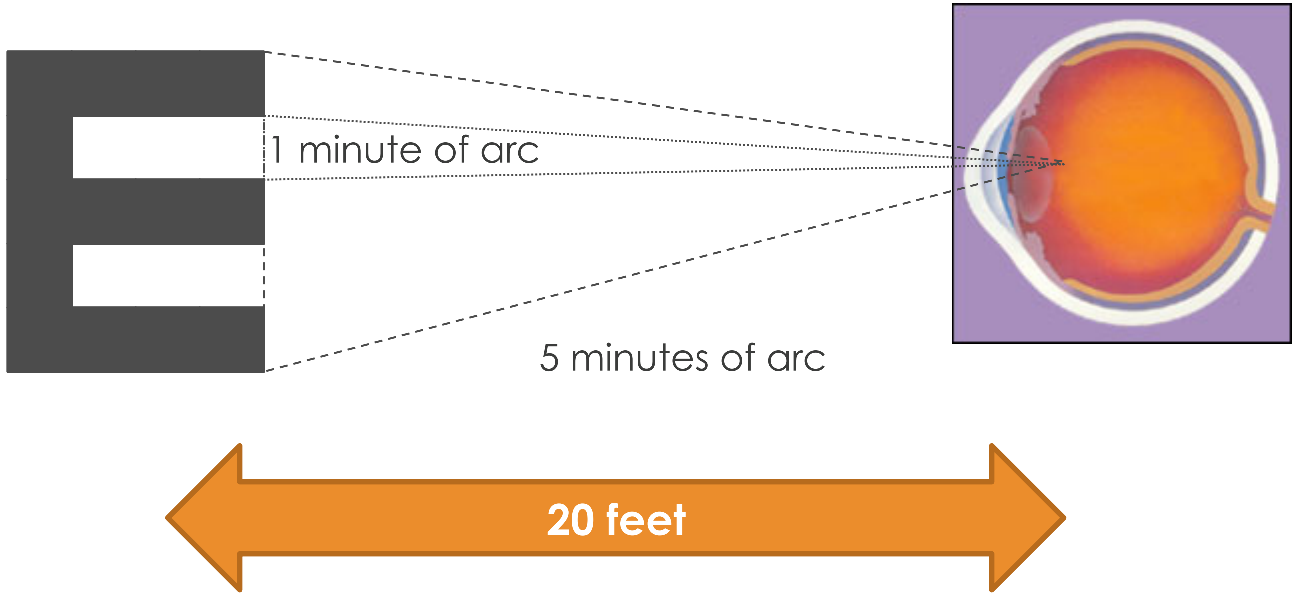
Simple Acuity → "Pixel Fusion Distance"
(~60 cycles per degree)

Snellen Acuity → "Reading"

Perceived Video Quality is a Function of Visual Acuity

Snellen Acuity

A person with 20/20 vision can resolve 1 minute of arc at 20 feet.



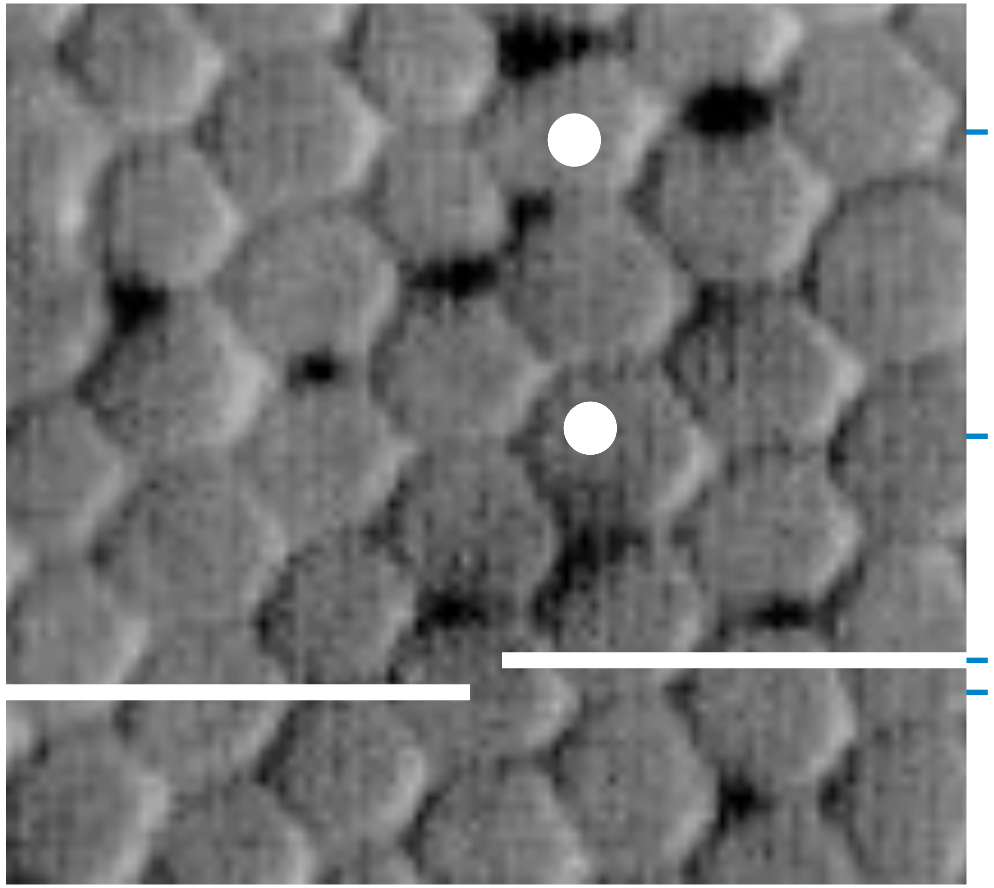
Equivalent to 30 cycles per degree of visual field

Snellen Acuity Matters

Rescaling Impacts Legibility

Hyperacuity

Prof Gerald Westheimer
UC Berkeley, Neurobiology & School of Optometry
Fellow, Royal Society of London
Order Australia
Fellow, American Academy of Arts and Science



Photoreceptor mosaic

Simple acuity

Result of photoreceptor sampling

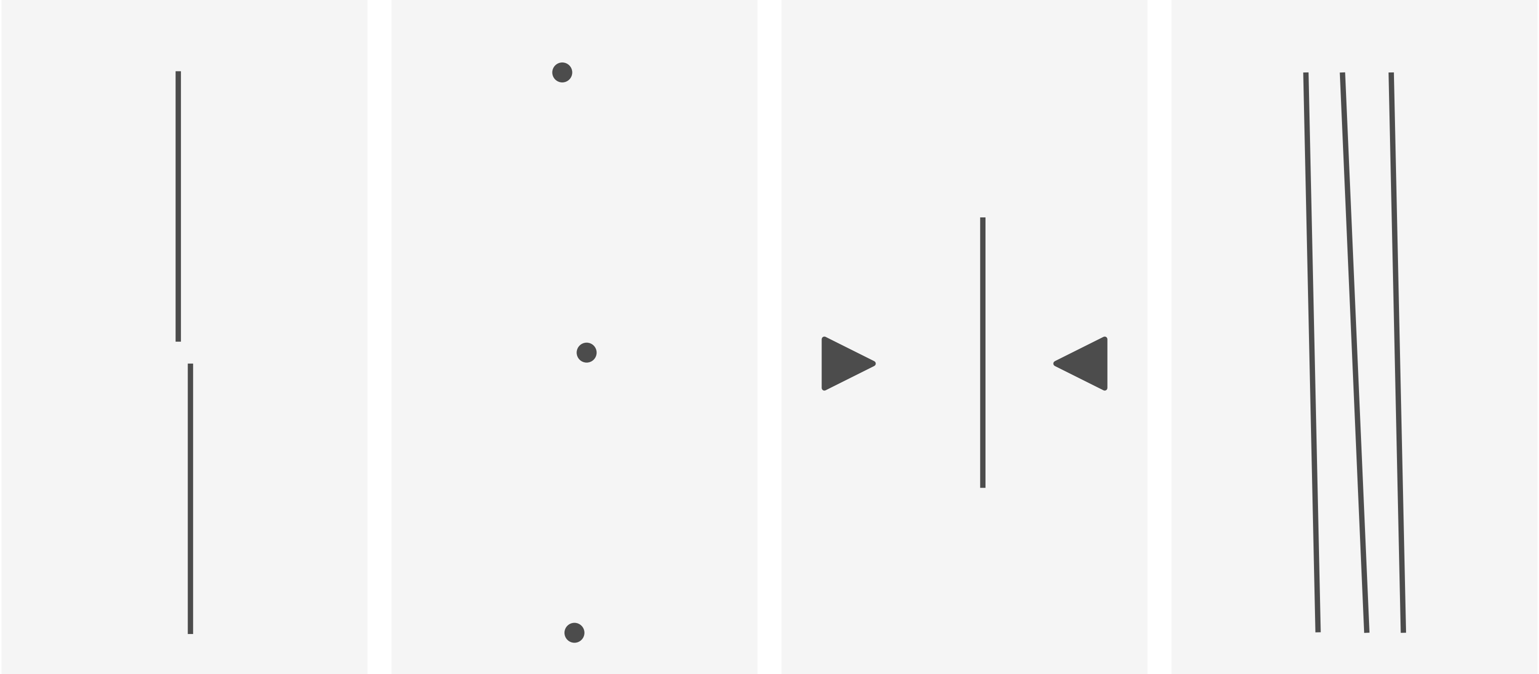
Limits data transfer into the visual system

Hyperacuity

Result of neural computation

Tells us about brain's internal model of the world

Hyperacuity



Vernier | Alignment | Separation | Tilt

Various Hyperacuity Tasks

~120 cycles per degree (depends on task)

Spatial Aliasing

“Stair Step” Artifact

fixed FOV – looking thru the window

HDTV	4K UHDTV	8K UHDTV	
32° FOV	**32° FOV**	**32° FOV**	
Simple Acuity	Hyperacuity	Hyperacuity	**display**
Snellen Acuity	Simple Acuity	Hyperacuity	**image**

SMPTE 2012

Smartphones & Tablets → "Windows"?

4.7" smartphone

16° FOV

qHD | **1080p**

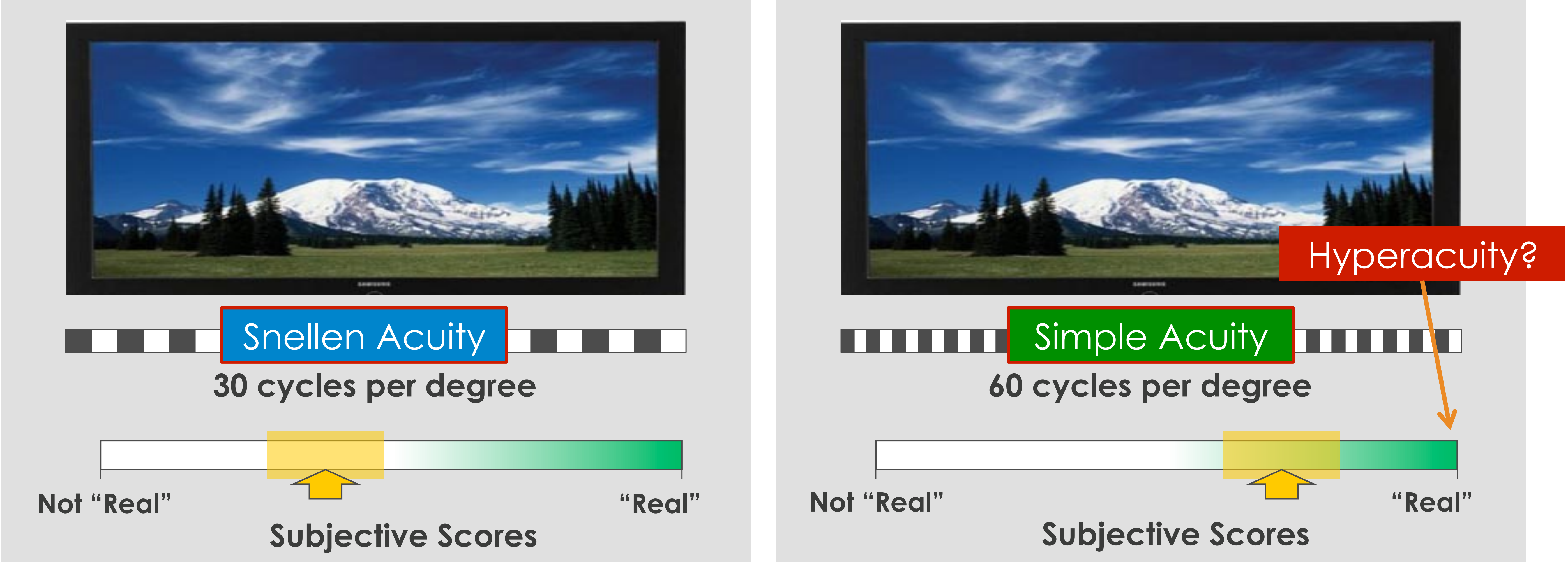
10" Tablet

32° FOV

1080p | **4K UHDTV**

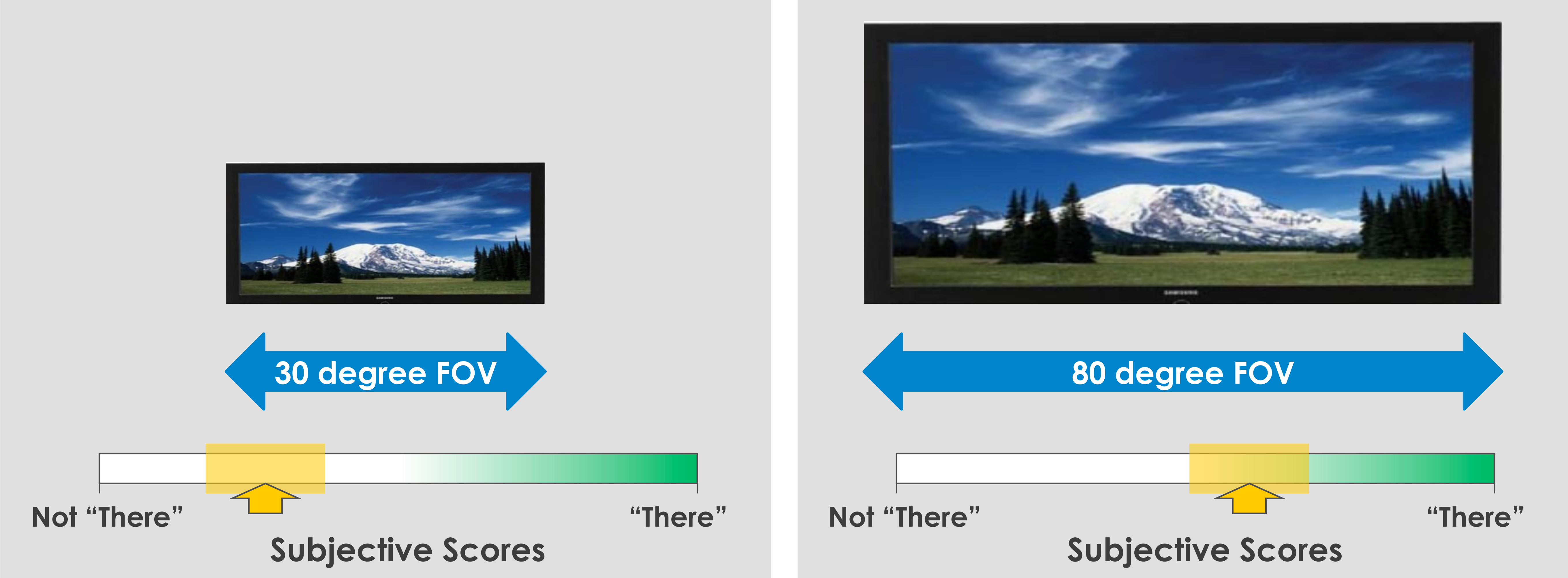
	4.7" smartphone: qHD	4.7" smartphone: 1080p	10" Tablet: 1080p	10" Tablet: 4K UHDTV
display	Simple Acuity	Hyperacuity	Simple Acuity	Hyperacuity
image	Snellen Acuity	Simple Acuity	Snellen Acuity	Simple Acuity

Sense of "Realness" Depends on Resolution



Based on research reported by Yamashita et al. "Super Hi-Vision" Video Parameters for Next-Generation Television. SMPTE Mot. Imag J. May-June 2012 vol. 121 no. 463-68 (NHK Science & Tech Res. Labs)

Sense of "Being There" Depends on Field of View



Based on research reported by Yamashita et al. "Super Hi-Vision" Video Parameters for Next-Generation Television. SMPTE Mot. Imag J. May-June 2012 vol. 121 no. 463-68 (NHK Science & Tech Res. Labs)

Wide Field of View

display	Simple Acuity
image	Snellen Acuity

~~Hyperacuity~~

qHD

16° FOV

HDTV

32° FOV

4K UHDTV

64° FOV

8K UHDTV

128° FOV

60 pixels per degree in all cases

Acuity, FOV, & Display Resolution

Type of Acuity	Acuity	Field of View (degrees)				
	cycles per degree	16	32	64	96	128
		Horizontal Resolution (pixels)				
Snellen	**30**	**480**	**960**	**1920**	**2880**	**3840**
Simple	**60**	**960**	**1920**	**3840**	**5760**	**7680**
Hyperacuity	**120**	**1920**	**3840**	**7680**	11520	15360
	180	**2880**	**5760**	11520	17280	23040
	360	**5760**	11520	23040	34560	460

UHDTV Supports Wide FOV

A sense of "Being There"

UHDTV Supports Hyperacuity

A sense of "Looking Through a Window"

Best of Both

SMPTE 2012

Mobile may be our first "windows"

Type of Acuity	Acuity cycles per degree	Field of View (degrees) 16	32	64	96	128
		Horizontal Resolution (pixels)				
Snellen	**30**	**480**	**960**	**1920**	**2880**	**3840**
Simple	**60**	**960**	**1920**	**3840**	**5760**	**7680**
Hyperacuity	**120**	**1920**	**3840**	**7680**	11520	15360
	180	**2880**	**5760**	11520	17280	23040
	360	**5760**	11520	23040	34560	46080

UHDTV Supports Wide FOV

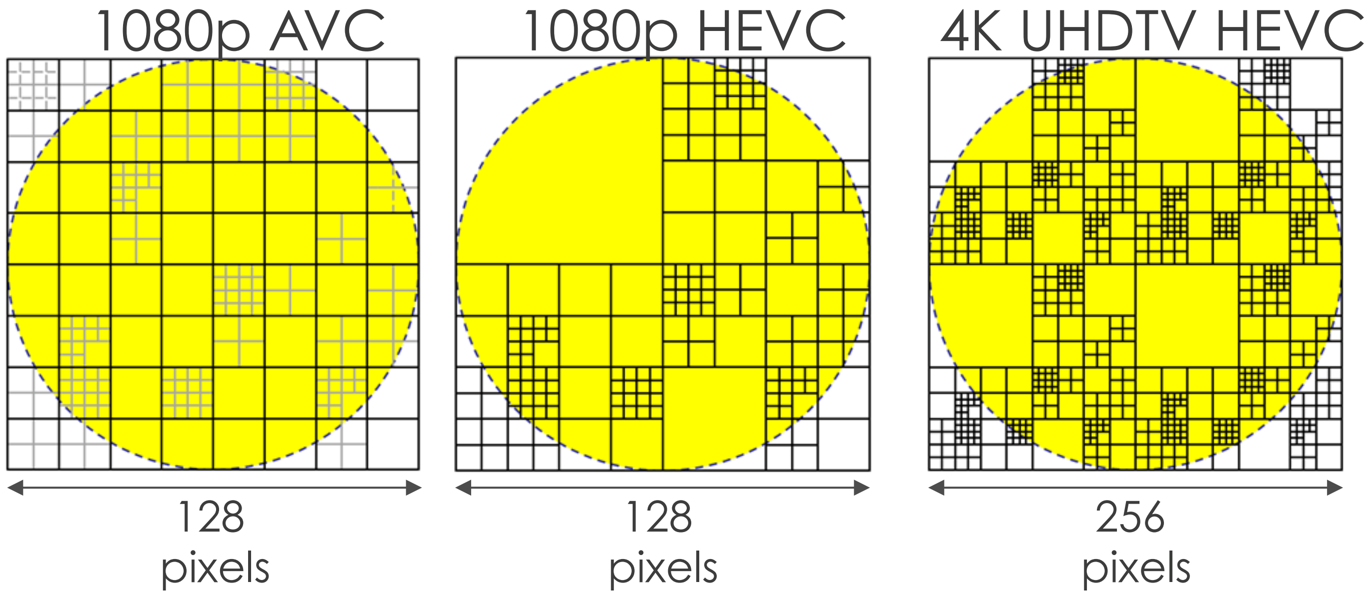
UHDTV Supports Hyperacuity

A sense of "Looking Through a Window"

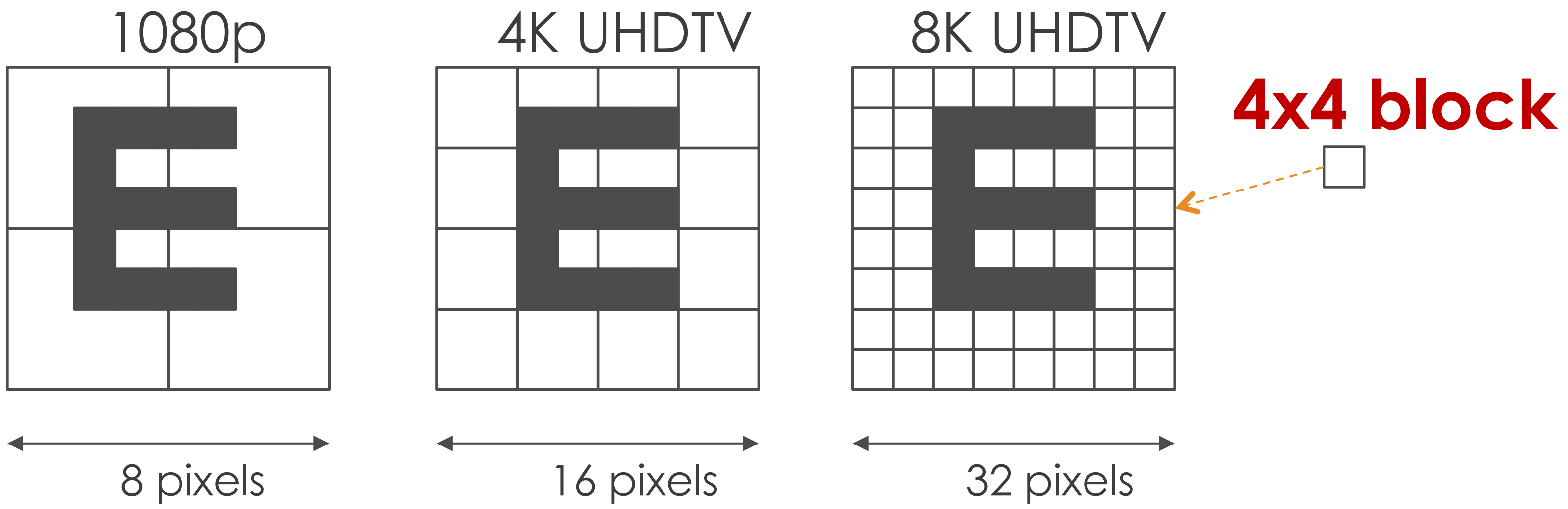
Can compression science keep up?

32° FOV

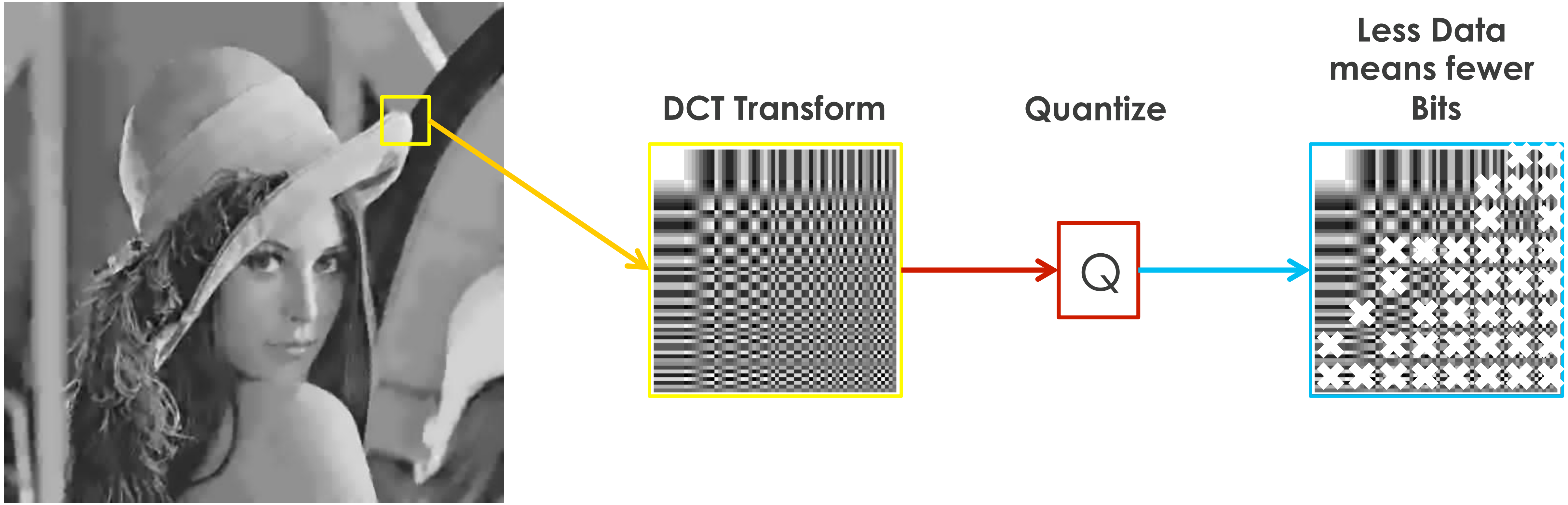
Coding partitions within the Foveola for 32° FOV



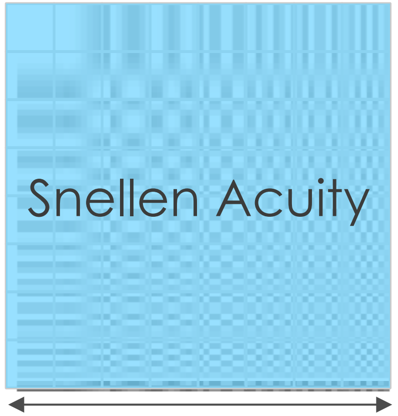
Coding partitions are large compared to the visual scale of even Snellen acuity



Compression



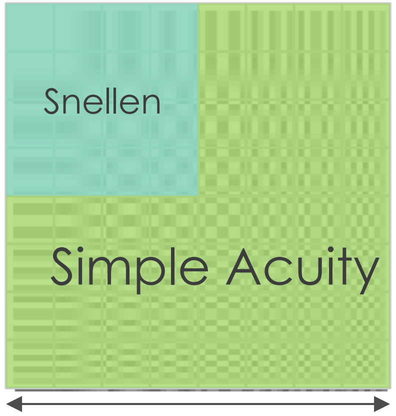
Acuity & DCT Subsets



HDTV

32° FOV

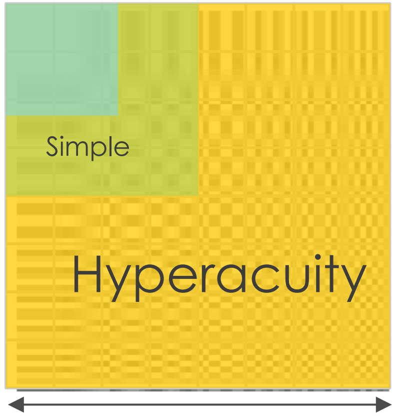
8 minutes of arc



4K UHDTV

32° FOV

4 minutes of arc



8K UHDTV

32° FOV

2 minutes of arc

DCT sub-domains map to progressively finer acuity thresholds

Smartphones & Tablets

4.7” smartphone

16° FOV

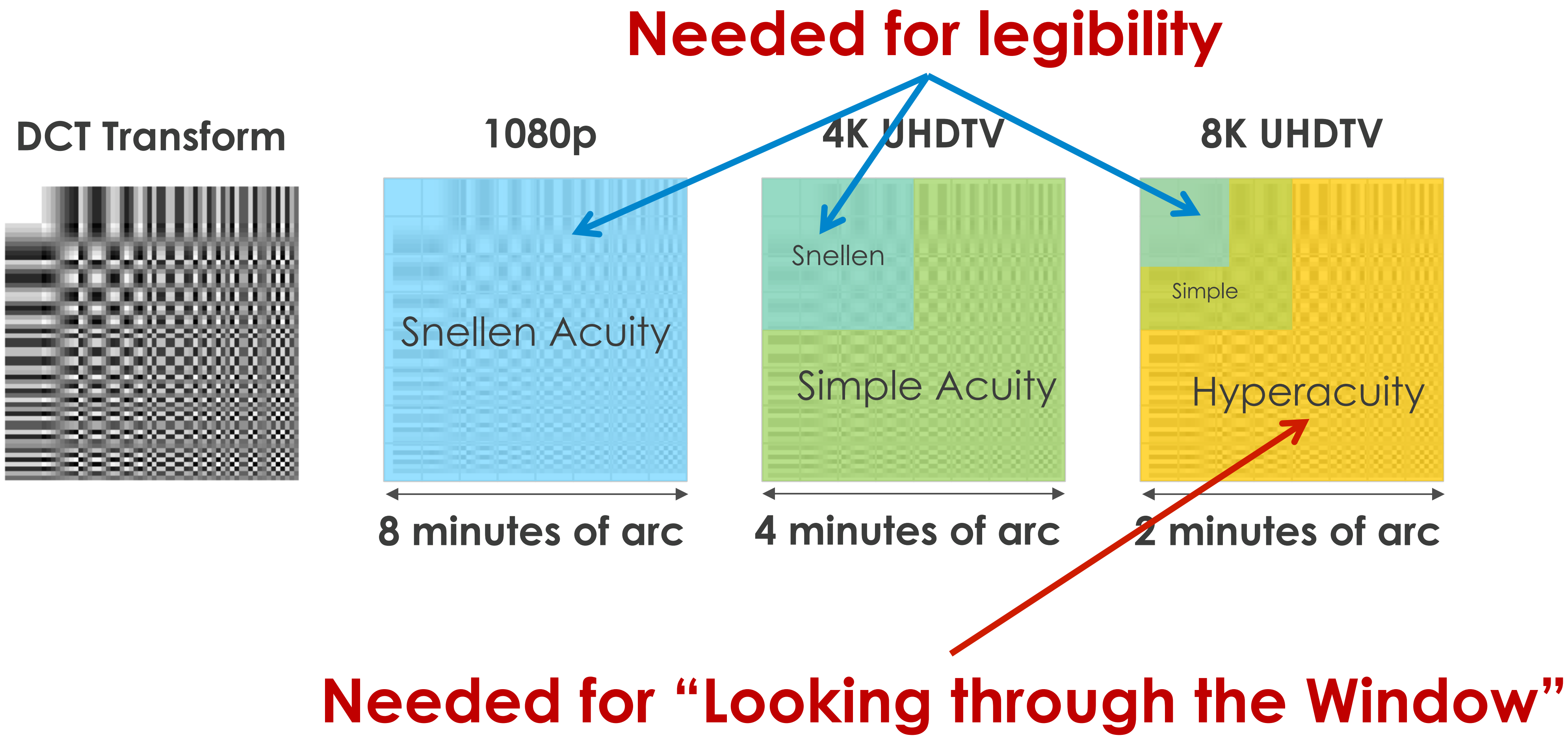
qHD	1080p
Snellen Acuity	Snellen / Simple Acuity
8 minutes of arc	4 minutes of arc

10” Tablet

32° FOV

1080p	4K UHDTV
Snellen Acuity	Snellen / Simple Acuity
8 minutes of arc	4 minutes of arc

Acuity & DCT Coefficients



Wide Field of View

qHD
16° FOV
Snellen Acuity
8 minutes of arc

HDTV
32° FOV
Snellen Acuity
8 minutes of arc

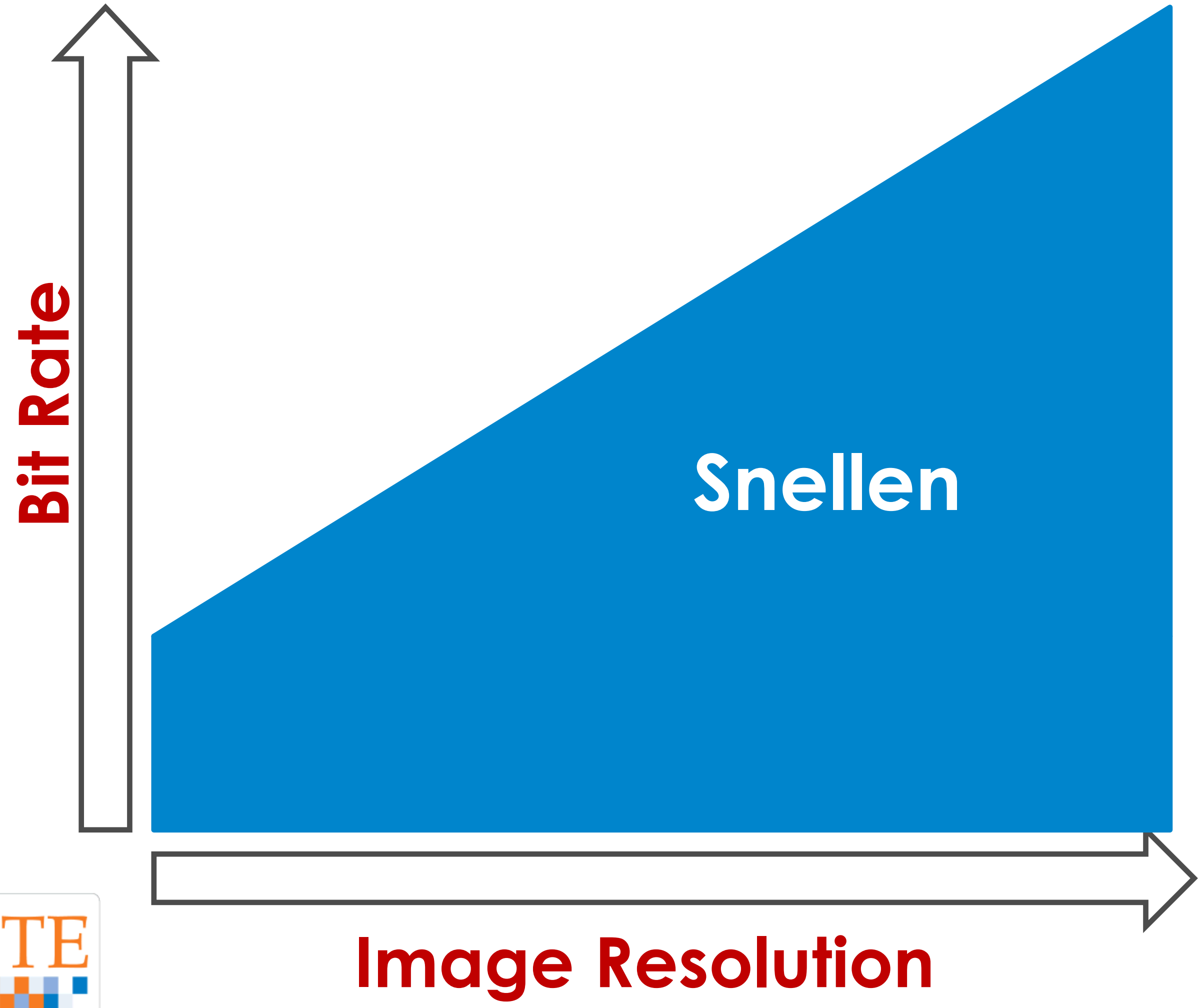
4K UHDTV
64° FOV
Snellen Acuity
8 minutes of arc

8K UHDTV
128° FOV
Snellen Acuity
8 minutes of arc

SMPTE 2012

Wide FOV and Bit Rate

Wide FOV



16° FOV 32° FOV 64° FOV 128° FOV

Image Resolution

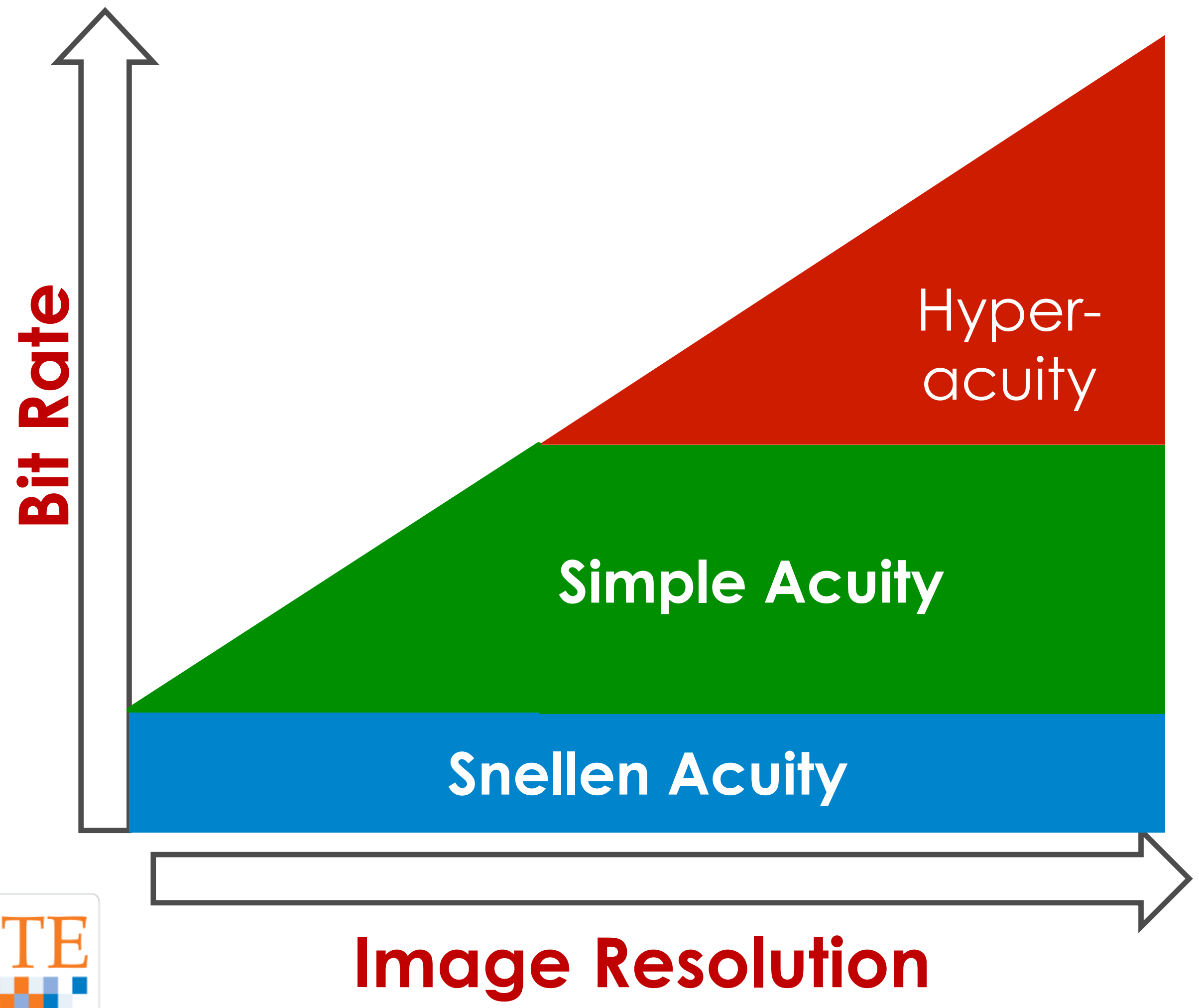
Bit rate should scale with Image Resolution.

~ 1080p Video Quality

SMPTE 2012

HEVC & UHDTV → Bit Rate Flexibility

32° FOV



HDTV
32° FOV

4K UHDTV
32° FOV

8K UHDTV
32° FOV

	AVC			HEVC		
	LO	MID	HI	LO	MID	HI
1080p	1	--	--	½	--	--
4K	1	4	--	½	2	--
8K	1	4	16	½	2	8

SMPTE 2012

Conclusions

Move toward higher resolutions in mobile and UHDTV leads to new competing viewing scenarios

"Being There" **Wide FOV**	**"Looking through the Window"** **Fixed FOV**

Mobile Devices may be the first "Windows"

Compression Efficiency will be more Important than Ever Before

HEVC & Beyond-HD → Better Bit Rate Flexibility

SMPTE 2012

Thank you

smccarthy@motorola.com

SMPTE 2012

SMPTE Meeting Presentation

Quantitative Evaluation of Human Visual Perception for Multiple Screens and Multiple CODECs

Sean McCarthy, Ph.D. Fellow of the Technical Staff

Motorola Mobility, 2450 Walsh Ave, Santa Clara, CA. smccarthy@motorola.com

Written for presentation at the

2012 SMPTE Annual Technical Conference & Exhibition

Abstract. *Great consumer experiences are created by a convergence of sight, sound, and story. This paper is an in-depth quantitative analysis of the neurobiology and optics of sight. More specifically, we examine how principals of vision science can be used to predict the bit rates and video quality needed to make video on everything from smartphones to Ultra HDTV a success. We present the psychophysical concepts of simple acuity, hyperacuity, and Snellen acuity to examine the visibility of compression artifacts for MPEG4/H.264. We also take a look at the newest emerging International compression standard HEVC. We investigate the how the various sizes of the new coding units in HEVC would be imaged on the retina, and what that could mean in terms of the HEVC video quality and bit rates we would likely need in order to deliver entertainment quality content to smartphones, tablets, HDTV, 4K TV, and Ultra HDTV.*

Keywords. Leave the word "Keywords." then type keywords or key phrases, separated by commas. List both specific and general terms that will aid in searches.

INTRODUCTION

This paper combines principles from vision science and compression science to understand better the visibility of video artifacts. The particular motivation behind writing this paper is the emergence of new ways of distributing media to consumers, and the new ways in which people watch video.

There is a clear trend lately of TV jumping from the living room onto smart phones, tablets, and laptops, a phenomenon abetted by new distribution models such as adaptive bit rate (ABR) streaming. At the same time, living room television displays are starting down a path towards higher resolutions beyond HDTV. 4K UHDTV displays have been announced[1] and the International Technical Union is developing industry Recommendations for both 4K and 8K UHDTV[2]. Making matters even more dynamic, a new way of compressing video is about to hit the market – the emerging High Efficiency Video Coding (HEVC) international standard.

This wealth of new technologies should prove to be a boon to consumers; but equipment manufacturers and service providers want to be sure they can continue to deliver great visual experiences that meet or exceed expectations. One challenge is that it is not clear how to quantify the perceptual response people might have to smaller and larger screens inside and outside the home.

The goal of this paper is to provide information and analysis to help predict and quantify viewer experiences using applied vision science. Use of vision science and models of the human visual systems to predict video quality is not new[3]. Rather, the aim of this paper is to expand that framework to include concepts such as hyperacuity – the real-world ability to discriminate visual features smaller than the size of a retinal photoreceptor – and Snellen acuity as a measure of the readability of text. We examine how the various kinds of visual acuity are likely to influence viewing distance and field of view. Finally we investigate the coding-block structure of AVC and HEVC to understand how the size of artifacts might impact their visibility on various displays.

VISUAL ACUITY

The retina is made of layers of specialized neurons each having a unique critical role in vision[4]. The light-gathering layer is the photoreceptor layer – a dense mosaic of cells that lines the eyeball and turn photons into neural signals.

The photoreceptor layer is constructed of different kinds of light-sensitive cells: red-, green-, and blue-sensitive cones; and achromatic rods. Cone photoreceptors dominate the central parts of the photoreceptor layer and are responsible for vision during the day. Rod photoreceptors dominate the periphery and are designed to take charge of seeing between dusk and dawn.

The topography of the retina is actually much more sophisticated than a blunt demarcation between center and periphery[5,6]. Figure 1 illustrates concentric zones of retinal topography and notes the maximum resolvable spatial-frequency for each zone. The tiniest center of the retina contains only red- and green-sensitive cones (no blue cones or rods) and is the center of our visual field. Polyak called this the central island of the foveola[7]. The central island corresponds to only approximately 0.2 degrees (12 minutes or arc) of the visual field, but it has the highest density of photoreceptors and is thus responsible for our most acute vision. The foveola, which contains the central island, spans approximately 1.2 degrees of visual angle. Being free of rods and blood vessels, the foveola also supports a very high cone density and high acuity vision. Encompassing the foveola is the fovea, which spans approximately 6 degrees of visual field. The fovea is composed of a mix of cones and rods with cones becoming scarcer and rods becoming denser moving away from the center. Moving outward from the fovea, the cones density continues to decline as does daylight acuity. Each eye has a blind spot approximately 15 degrees to the side of the fovea. (The blind spot is the photoreceptor-free optic disk created by the optic nerve as it passes through the photoreceptor layer on the way to the brain.) Together, the right and left blind spot span about 30 degrees and would tend to flank the edges of an HDTV screen at normal viewing distances.

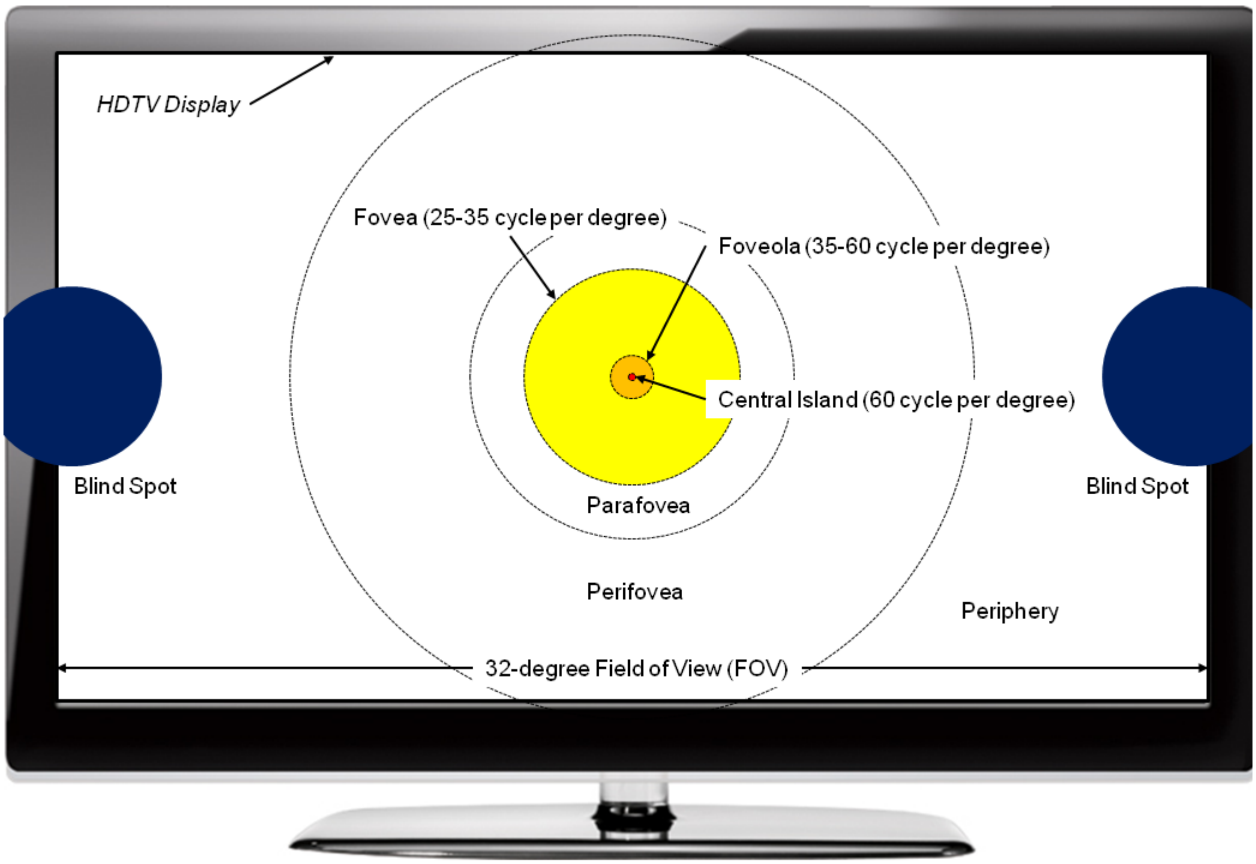


Figure 1. The photoreceptor layer as projected onto visual space.

Cone photoreceptors in the central island of the foveola are arranged in a close-packed hexagonal array. The spacing of cones is approximately ~30 seconds of visual field (~2 cones per minute of arc), which translates to ~120 cone photoreceptors per degree. According to Nyquist sampling theory, the maximum theoretically achievable resolving power of the retina would thus be about 60 cycles per degree (1/2 the number of sampling points). It turns out that the eye's maximum resolving power matches the mathematical limit[8] – 60 cycles per degree, at least under ideal conditions. (In the real world, the maximum theoretical acuity can be hard to match for several reasons[9]. Moreover, some people need glasses, others have slightly clouded corneas, and sometimes the ambient lighting or image contrast is just not right.)

For 150 years, we've been able to say that a person with normal vision has 20/20 vision. Herman Snellen[10] gave us this way of talking about visual acuity in 1862. The term "20/20" is a Snellen fraction that means that a person with normal acuity can read or identify a particular sized stylized letter or symbol – called an optotype -- from 20 feet. (In meters, 6/6 is used for normal acuity.) Snellen-style optotypes are designed so that each symbol on the "20/20" row would subtend 5 minutes of arc, and each gap would subtend 1 minute of arc when viewed from 20 feet, as illustrated in Figure 2.

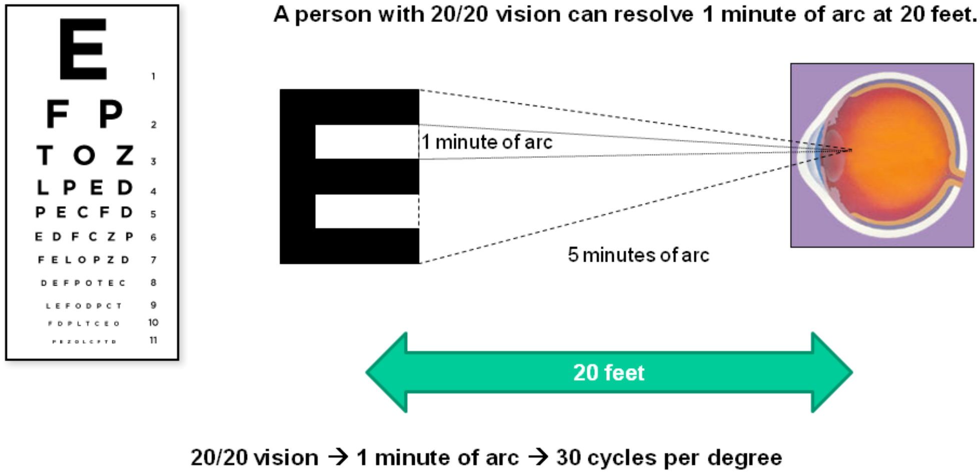


Figure 2 A Snellen chart is made up of rows of optotypes (engineered letters or symbols) of monotonically increasing size. The optotype "E" shown has 5 horizontal segments, which is equivalent to 2.5 cycles per 5 minutes of arc – or 30 cycles per degree.

One might expect that 20/20 vision would be equivalent to the Nyquist-limited simple acuity of 60 cycles per degree. Instead, 20/20 vision corresponds to acuity of 30 cycles per degree as explained in Fig. 2. Snellen-style optotypes are used to measure "reading acuity" – or the ability to identify structured symbols – for "normal" people under "normal" viewing conditions, whereas simple acuity relates a theoretical optical limit that can only be achieved by the keenest of eyes under exceptional viewing conditions.

In 1975 Westheimer coined the term hyperacuity[11] to put a handle on a person's ability to discriminate the location of visual features on a scale far finer than the limit imposed by the Nyquist limit. Indeed, as remarkable as it may seem, we routinely notice visual features smaller than the diameter of a cone photoreceptor[12]. Examples of hyperacuity tasks are shown in Figure 3. In video compression, the kinds of features that are likely to invoke hyperacuity are spatial aliasing -- "stairstep artifacts" -- that can be introduced by rescaling, ringing and mosquito noise around letters and other hard edges, slight mismatches during motion estimation, de-interlacing artifacts, and more.

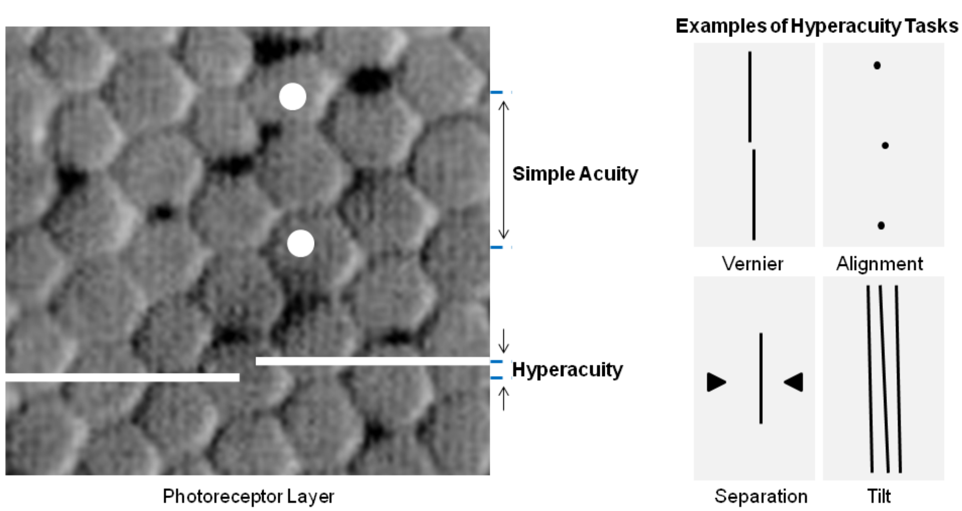


Figure 3. Illustration of photoreceptor mosaic and the retinal image of stimuli that correspond to simple (Nyquist) acuity (2 dots) and hyperacuity (misaligned line segments). Vernier acuity is one of the easiest forms of hyperacuity to appreciate, but we also experience other forms of hyperacuity every day. Hyperacuity relates our ability to localize features.

Hyperacuity arises from a combination of optics and neurobiology. The optical limits imposed on simple acuity also limit hyperacuity, but hyperacuity arises because groups of neurons work as a collective to process light stimuli. As such, hyperacuity tells us something about the brain's internal model of the world, whereas simple acuity tells us something about the limits of data transfer into the visual system.

VISIBILTY OF ARTIFACTS

The various forms of visual acuity allow us to see both display artifacts and compression artifacts. Display artifacts arise because the density of the pixels imaged onto the retina is insufficient to mask the visibility of individual pixels (seen with simple acuity) or distortions caused by groups of pixels (seen with hyperacuity). If one steps away from the display, our ability to discriminate pixels or notice tiny distortions diminishes. At some point, the imaged pixel density would be so high as to be beyond our hyperacuity threshold: The display would take on the characteristics of a window. Any artifacts we could see in such a situation would be due to compression artifacts or actual features in the original image.

Table 1. Relationship between Acuity, Field of View, and Display Resolution

Type of Acuity	Acuity Threshold: seconds of arc	Acuity Threshold: cycles per degree	Field of View (degrees): 16	32	64	96	128	
			Horizontal Resolution (pixels)					
Reading Acuity	60	30	480	960	1920	2880	3840	UHDTV Supports Wide FOV
Simple Acuity	30	60	960	1920	3840	5760	7680	
Hyperacuity	15	120	1920	3840	7680	11520	15360	
	10	180	2880	5760	11520	17280	23040	
	5	360	5760	11520	23040	34560	46080	

UHDTV Supports Hyperacuity

Viewing distance and field of view (FOV) (the visual angle subtended by a display) are two sides of the same coin. A small smartphone held near the eye can subtend the same visual angle as an enormous flat panel display across the room. What is important in determining the vantage point at which pixels disappear from sight is the number of pixels per visual degree.

Table 1 provides insight into today's smartphones, tablets, and HDTV displays. The blue- and green- shaded cells in the upper left highlight typical FOV and display resolutions for smartphones. A 4" wide smartphone held at about 14" (16-degree FOV) would need a horizontal display resolution of at least 480 pixels to support "reading" acuity, and individual pixels would be visible even for a 960-pixel wide display. True 1080p screens would be required to make pixels and display artifacts disappear. The green-shaded cells are representative of tablets and HDTV's in the home. A 10" tablet held at a 17" viewing distance results in the same

32-degree FOV as a 48” wide HDTV viewed from 7’. In both cases, a horizontal resolution of 1920 pixels is sufficient to support simple acuity (masking individual pixels) but not hyperacuity (groups of pixels can still create distortions).

The emergence of beyond-HDTV resolutions on tablets, laptops, and home television displays provides two distinct viewing scenarios. First, UHDTV (4K & 8K) enables a much wider FOV. Wide FOV has been shown to promote a subjective sense of realness and “being there”[13] But note that even for 8K UHDTV, the wide FOV results in display performance only at the simple acuity level. In the second viewing scenario, the FOV is about the same as we are used to today when viewing a tablet or home HDTV display, but UHDTV crosses the hyperacuity threshold – watching a UHDTV would become like looking out a window, or at least the next best thing. In such situations, video quality would be independent of the display-induced distortions and would thus be limited only by the quality of video compression.

Figure 4 illustrates the size of AVC[14] and HEVC[15,16] basic coding partitions relative to the central 1-degree of visual field approximately corresponding the to foveola. Note that the smallest coding partition for both AVC and HEVC is 4 pixels by 4 pixels, and thus are the same size when imaged on the retina. Also shown are the retinal images of an “E” optotype for the same viewing conditions.

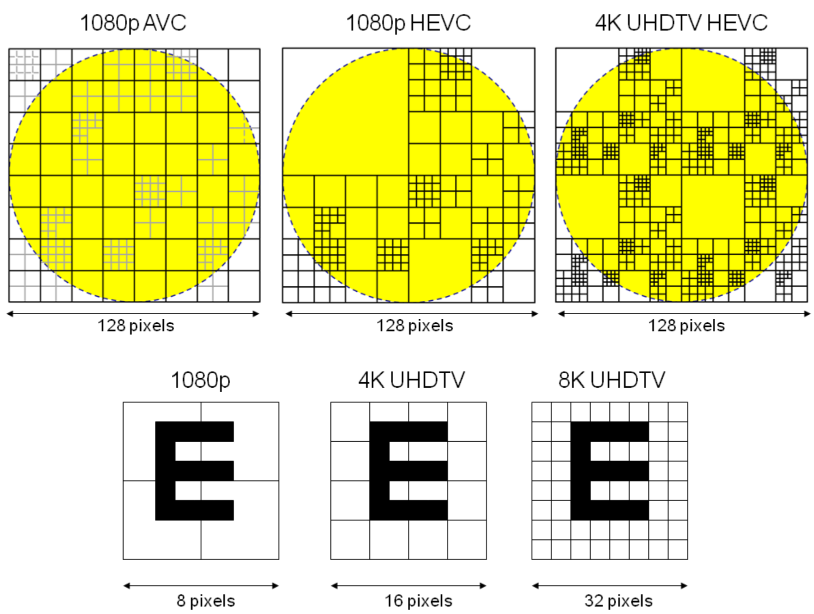


Figure 4. Illustration of size of AVC macroblocks and subpartitions and HEVC Coding Units relative to the central 1-degree of visual field (yellow inscribed circle) for 1080p tablets and HDTV displays subtending a 32-degree field of view, and HEVC coding units for a 4K UHDTV Untitled eventalso subtending a 32-degree FOV. Lower row: the retinal image of an “E” optotype for 1080p, 4K, and 8K displays for 32-degree FOV. Each of the smallest blocks for each display type represents a 4x4 coding partition.

Note in Figure 4 that 8K UHDTV (32-degree FOV) is required before the coding partition size matches the scale of Snellen optotypes. While far superior to today’s 1080p resolutions, the size of HEVC coding units are still large compared to the retinal scale of simple acuity and hyperacuity tasks. The implication is that the size of coding units – not the resolution of the display – will be a limiting factor with regard to video quality, even when HEVC becomes ubiquitous.

One of the most important techniques in modern video compression is the use of spatial-frequency decomposition. AVC and HEVC both use a decomposition method called Discrete Cosine Transform (DCT), and HEVC also uses a related phase-shifted method called Discrete Sine Transform (DST). AVC is limited to 4x4 and 8x8 transforms, but HEVC can also use 32x32 and 16x16 versions. Compression is achieved in part by eliminating or reducing the numeric precision of the coefficients used to code the various spatial frequency components. The general idea is to eliminate coefficients that would not produce noticeable distortions.

As illustrated in Figure 5, DCT transforms map coefficients from low to high spatial frequency moving in a zigzag manner from the upper left corner. The spatial-frequency of the lower right corner is set by the Nyquist limit of the display (which is different than the Nyquist limit of the photoreceptor layer, see Figure 6.)

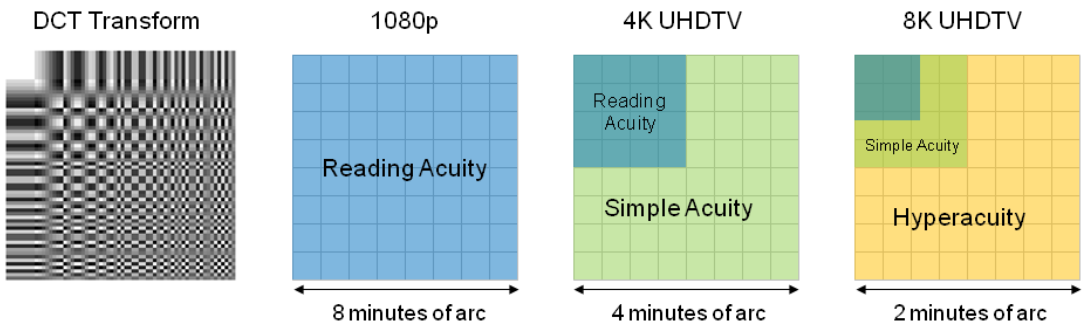


Figure 5. Illustration of the progressive shift of DCT coefficients into finer grains of acuity with increasing display resolution. The scenario depicted corresponds to a FOV or 32 degrees and represents an 8x8 DCT.

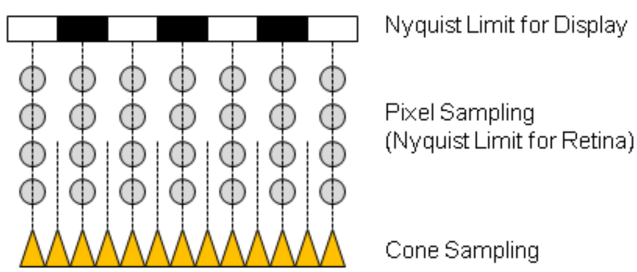


Figure 6. Nyquist limit of the retinal image of a picture on a display is typically ½ that of the photoreceptor layer. The photoreceptor layer sees the entire display – pixels and spaces between pixels, which together constitute a single effective cycle. On the other hand, the display is limited to one cycle per 2 pixels, which is equivalent to approximately 4 photoreceptors when the display is held at a distance that corresponds to the simple acuity limit.

By associating the DCT coefficients with the various acuity thresholds, it becomes clearer just which DCT coefficients might have the biggest impact for different display types and video-quality targets. In moving from a 1080p display to a 4K display, for example, Figure 5 and Table 2 indicate that spatial-frequency coefficients in the upper-left corner would provide the level of image detail need for Snellen acuity. The remainder of the coefficients would code image details that would be visible using a simple-acuity threshold criterion. Similarly for 8K UHDTV, the upper-left region corresponds to image details – and artifacts – at the Snellen-acuity and simple-acuity levels, whereas the bulk of the coefficients would code image details on a scale visible only at the hyperacuity level.

Table 2. Comparison of Display Nyquist Limit and Various Forms of Visual Acuity

Device Type	Horizontal Resolution	Field of View (degrees)				
		16	32	64	96	128
		Display Nyquist Frequency (cycles per visual degree of arc)				
Smartphone	960	30	15	8	5	4
Tablets & HDTV	1920	60	30	15	10	8
4K UHDTV	3840	120	60	30	20	15
8K UHDTV	7680	240	120	60	40	30

Table 2 provides a more quantitative adjunct to Figure 5. The right-hand columns list the Nyquist limit of displays having progressively higher resolutions. Each cell is color coded to indicate the level of corresponding visual acuity – reading acuity (blue), simple acuity (green), and hyperacuity (orange). For example, a 1080p tablet subtending 32-degree FOV would support display of spatial frequencies only up to the "reading" acuity level. But a 4K UHDTV subtending the same FOV would support display of spatial frequency at the simple acuity level. A similarly viewed 8K UHDTV would support display of details into the range of hyperacuity.

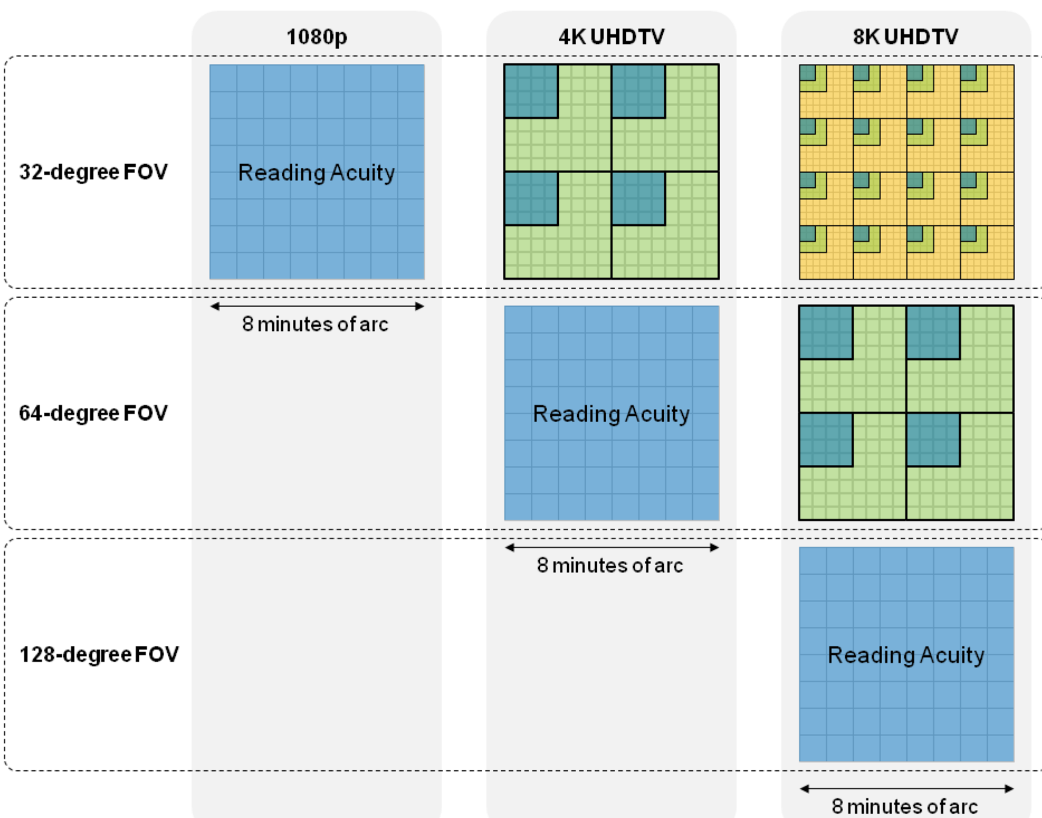


Figure 7. Illustration of the different importance of DCT sub-domains for various viewing conditions. Each grid represents 8 degrees of the retinal image, which corresponds to one or more 8-by-8 DCT transform matrices.

One of the objectives of this paper is to develop a better idea of the bit rates that are likely to be needed to meet or exceed consumer expectations. As shown in Figure 7, the importance of various sub-regions of the DCT coefficients can be mapped to viewing angle (and thus viewing distance) and display resolution in terms of various levels of visual acuity. In general, the fewer DCT coefficients that need to be coded the lower the bit rate needed.

An implication of Figure 7 is that the most inflexible compression environment corresponds to viewing environments that support only the lowest acuity tasks. These are the wide FOV scenarios in which each 8-by-8 DCT matrix spans the 8 minutes of arc of the visual field for 1080p, 4K, and 8K. In these situations, all of the DCT coefficients go toward coding visible detail at the level of Snellen acuity, we should thus expect that bit rate requirements would scale with the number of total pixels – 4-fold for 4K and 16-fold for 8k compared to 1080p.

An alternative message from Figure 7 is that 4K and 8K should provide a much greater degree of bit rate flexibility. Note that the level of acuity – a proxy for video quality – that can be achieved for a 1080p display having a 32-degree FOV can also be achieved by a fraction of the DCT coefficients needed for 4K and 8K 32-degree field of views. Indeed, the number of non-zero coefficient might be expected to be about the same in each case. Thus, we might expect that the minimum bit rate needed for 4K and 8K to support the video quality at the level we see today for HDTV should be about the same as we use today. Higher bit rates would tend to go towards spatial-frequency coefficients that would support higher levels of video acuity and hence video quality.

CONCLUSIONS

The objective of this paper has been to move towards putting intuitively true statements such as “higher resolutions and screen size will produce better viewing experiences” on a more quantitative basis. It is also the author’s hope that this paper will help foster the application of the principles of vision science to the needs of video compression and processing.

One of the key takeaways from this paper is that the move towards UHDTV displays can have 2 different -- and competing – benefits. The first is that UHDTV supports a wide field-of-view that people may find more real and immersive. The second benefit is that UHDTV is capable of turning televisions into the next best thing to a window in which neither pixels nor compression artifacts are visible, but only for a FOV similar to what most people experience today on 1080p tablets and home HDTVs. Even 8K UHDTV cannot simultaneously support immersive FOV and hide compression artifacts beyond our higher acuity thresholds.

Based on the size of coding partitions, as imaged onto the retina, it is likely that compression efficiency will continue to be the limiting factor with regard to video quality. Even in HEVC, and even for narrow-FOV 8K UHDTV, the minimum size of coding units is at least as large as the smallest feature on a Snellen optotype. Fortunately, HEVC is emerging from the International Standards development process in time to provide the efficiency boost needed. HEVC is on track to be twice as efficient as AVC: The same quality at half the bit rate, or twice the quality at the same bit rate – with the potential to deliver even better video quality at the same bit rate for narrow-FOV UHDTV scenarios.

A significant and potential cost-saving emergent property of 4K and 8K UHDTV is that they are likely to provide more opportunities for bandwidth management. We can reasonably expect that the status quo of HDTV video quality will be achievable at bit rates not much greater than are used today for HDTV service, or about half that for HEVC. We will also be able to improve video quality by smart use of the spatial-frequency coefficients that correspond to higher levels of visual acuity. The range of bit rates between baseline service and improved service should provide a useful degree-of-freedom for optimizing services to individual subscribers, and for managing groups of services through statistical multiplexing, for example. It may also facilitate the staged roll-out of beyond HDTV services in which incremental levels of video quality are layered onto services as HEVC and UHDTV technologies become more widespread.

Perhaps, one of the biggest unknowns at this point is how consumers will decide to view next-generation displays. Will they immerse themselves in video that covers most of their visual space, or will they stay with a narrower field of view that is familiar but gain a "through the window" level of video quality. In both cases, HEVC is likely to take a central role in delivering the bits more efficiently, though how we tune our compression algorithms will depend more than ever on how consumers hold their tablets and where they sit in the living room.

REFERENCES

1. "CES 2012: 4K TV sets make their debut, minus the hoopla," Los Angeles Times, January 11, 2012
2. ITU-R Study Group 6. "Ultra High Definition Television: Threshold of a New Age." Geneva, May 24, 2012.
3. S. McCarthy. " A Biological Framework for Perceptual Video Processing and Compression," SMPTE Mot. Imag. J; 119:(8) 24-32, November 2012
4. H. Kolb. "Simple Anatomy of the Retina." Webvision:The Organization of the Retina and Visual System. http://webvision.med.utah.edu/ 2012
5. H. Kolb et al. "Part XIII: Facts and Figures concerning the human retina." Webvision:The Organization of the Retina and Visual System. http://webvision.med.utah.edu/ 2012
6. C.A. Curcio et al. "Human photoreceptor topography". J Comp Neurol. 292:497–523. 1990
7. S.L. Polyak. *The Retina.* The University of Chicago Press: Chicago, 1941.
8. D.G. Green DG. "Regional variations in the visual acuity for interference fringes on the retina". J Physiol. 207:351–356, 1970
9. M. Kalloniatis and C Luu. "Visual Acuity" Webvision:The Organization of the Retina and Visual System. http://webvision.med.utah.edu/ 2012
10. H. Snellen. *Probebuchstaben zur Bestimmung der Sehscharfe (trans. Sample Letters for Determination of Visual Acuity)*. Utrecht: PW van de Weijer, 1862.
11. G. Westheimer. "Editorial: Visual acuity and hyperacuity". Investigative Ophthalmology 14 (8): 570–2. 1975
12. G. Westheimer. "Hyperacuity," Scholarpedia, 6(8):9973, 2011
13. T. Yamashita et al. "Super Hi-Vision" Video Parameters for Next-Generation Television," SMPTE Mot. Imag J. ; 121:(4) 63-68, May-June 2012
14. T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," IEEE Trans. Circuits Syst. Video Technol., vol. 13, no. 7, pp. 560-576, July 2003.
15. M.S. Goldman. "High-Efficiency Video Coding (HEVC): The Next Generation Compression Technology," SMPTE Mot. Imag. J; 121:(5) 34-39, July-August 2012
16. ISO/IEC JCT1/SC29/WG11 (MPEG), "Description of High Efficiency Video Coding (HEVC)," doc. no. N11822, Daegu, KR, January 2011.