Springer Series in Advanced Microelectronics 26

High-Dynamic-Range (HDR) Vision

Microelectronics, Image Processing, Computer Graphics

Bearbeitet von
Bernd Hoefflinger

1. Auflage 2007. Buch. xvi, 241 S. Hardcover
ISBN 978 3 540 44432 9
Format (B x L): 15,5 x 23,5 cm
Gewicht: 603 g

1

The Eye and High-Dynamic-Range Vision

Bernd Hoefflinger

The dream of electronic vision is to mimic the capabilities of the human eye and possibly to go beyond in certain aspects. The eye and human vision overall are so versatile and powerful that any strategy to realise similar features with electronics and information technology have to be focused on certain aspects of man's powerful sense to collect the most comprehensive information on our physical world.

The single most challenging feature of optical information is the high dynamic range of intensities hitting the same receptor from one instant to another or hitting different receptors in an observed scene at the same time. This range may exceed 8 orders of magnitude (Table 1.1).

This is wider than our eyes can handle instantaneously or even with long-time adaptation. The sensing and acquisition system in our eyes handles over 5 orders of magnitude in real time (Fig. 1.1) and up to 8 orders with long-time adaptation. The curve shown here represents the instantaneous response of a normal eye. The minimum detectable signal moves up and the maximum tolerable because of overflow (blinding) and damage decreases with age so that the dynamic range deteriorates with age. In any event, the eye's dynamic range used to be much superior to any real-time electronic acquisition system until the advent of certain high-dynamic-range (HDR) video sensors and cameras.

It is the cornerstone of this book that real-time HDR video acquisition with a dynamic range of more than 7 orders of magnitude has evolved to a mature technology since it was first demonstrated in 1992 [1]. Fundamentally, the pixels in these sensors mimic the logarithmic compression of the high input range similar to the photoreceptors in our eyes making them "bionic" sensors in many respects like contrast sensitivity, separation of the illuminant and color constancy [2]. Cameras with such sensors have been available since 1996 and a significant amount of research and development along the vision chain including video processing and the challenging art of displaying HDR video to our eyes has been performed.

Table 1.1. Optical intensities in photometric units in some real-world scenes

Condition	Illuminance (lx)
Clear night sky	0.001
Quarter moon	0.01
Full moon	0.1
Late twilight	1
Twilight	10
Heavy overcast	100
Overcast sky	1,000
Full daylight	10,000
Direct sunlight	100,000

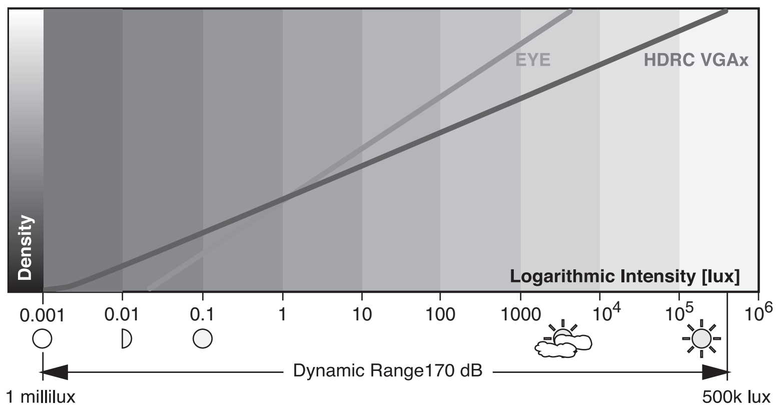

Fig. 1.1. The response, in the following also called the optoelectronic conversion function (OECF), of the human eye and of an HDRC array sensor

As a result, a comprehensive and exemplary treatment of HDR video can be presented, which, as we shall see, is inspired in many ways by the unique capabilities of the human eye and the human visual system (HVS).

The eye-like HDR video acquisition paradigm in this book is evident in Fig. 1.1, where the OECF of a commercial High-Dynamic-Range CMOS (HDRC) sensor is overlayed on the eye's OECF.

What is striking is that the electronic sensor is superior to the eye in:

1. Sensitivity, which means the minimum detectable illuminance, by at least a factor of 10
2. Range, which means the maximum illuminance measurable without distortion, white saturation or damage by another factor of 10, so that the total dynamic range is more than 100× higher than that of the human eye

The unique feature of this sensing technology is HDR photometric video capture in real time of 8 orders of magnitude of illuminance.

This means that over an illuminance range of 100 million to 1 all values measured consecutively on the same pixel and/or simultaneously on different pixels in an array are correct relative to each other and on an absolute scale, if the sensor temperature is kept constant and the sensor is calibrated.

Before we treat other unique features and consequences of this paradigm of real-time photometric HDR video capture, it is useful to put our treatment of HDR vision in perspective relative to the book HDR IMAGING by Reinhard, Ward, Pattanaik and Debevec [3], which we will reference as HDRI in the present book (HDRV).

HDRI is a formidable scientific treatment on the "processing" of high-dynamic-range images acquired with low-dynamic-range cameras basically through multiple exposures. This means HVI is focused on high-quality still images and computer graphics while there are limitations for video, particularly high-speed video in real-world scenes.

The present book HDRV presents the disruptive technology of HDR "acquisition" with eye-like log-compression up front in the electronic photoreceptor (pixel) and its powerful consequences for robust, real-time video processing, transmission and display. Our focus on real-time HDR video capture is demonstrated in Figs. 1.2–1.4.

Fire breather, solar eclipse and welding simply could not be captured without the high-speed, real-time recording offered by the HDRC sensors. However, even with the "natural", eye-like log compression in each pixel, the camera output data for these scenes still needed "tone mapping", which will be presented

Fig. 1.2. A frame taken from the video sequence of a fire breather recorded with an HDRC video camera at 30 frames/s^{-1}

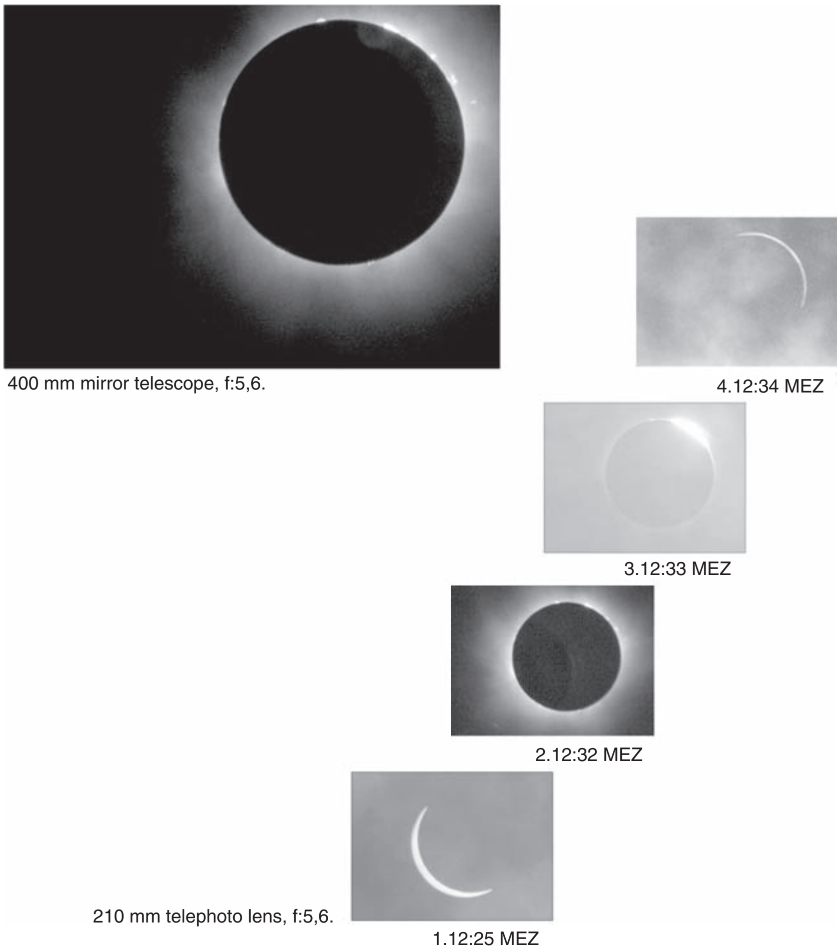

Fig. 1.3. Solar eclipse: a few frames from the video sequence of the solar eclipse on August 11, 1999, in Germany, recorded with a HDRC camera at 30 frames/s^{-1}. The camera was equipped with a 400 mm telescope, f:5.6, and aH6 filter. The partially clouded dark sky was genuinely captured with its shady details because of the low-light contrast sensitivity of the HDRC sensor simultaneously with the extremely bright corona

in later chapters, to offer us an informative print image. It is in areas like tone mapping where all in the HDR community come together to benefit from the general progress in HDR imaging.

However, we return here to HDR eye-like logarithmic video acquisition to identify further tremendous benefits resulting from mimicking the HVS.

Figure 1.5 shows HDR acquisition. The scale in this graph is powers of 2, also called f-stops or octaves. The upper bars show how seven exposures

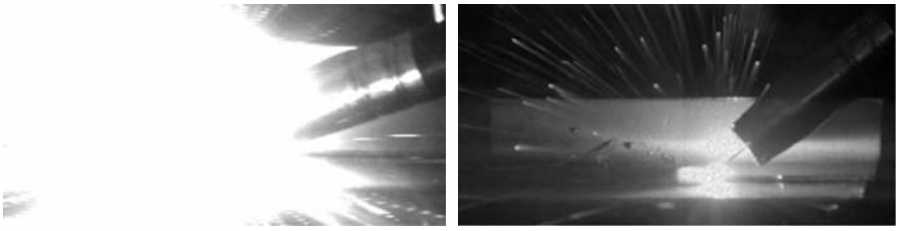

Fig. 1.4. Welding scence: Single frames from the video sequence of a welding scene recorded simultaneously with a CCD and an HDRC camera in 1995. The HDRC sensor had a resolution of 256×128 pixels and it delivered 120 frames/s^{-1}

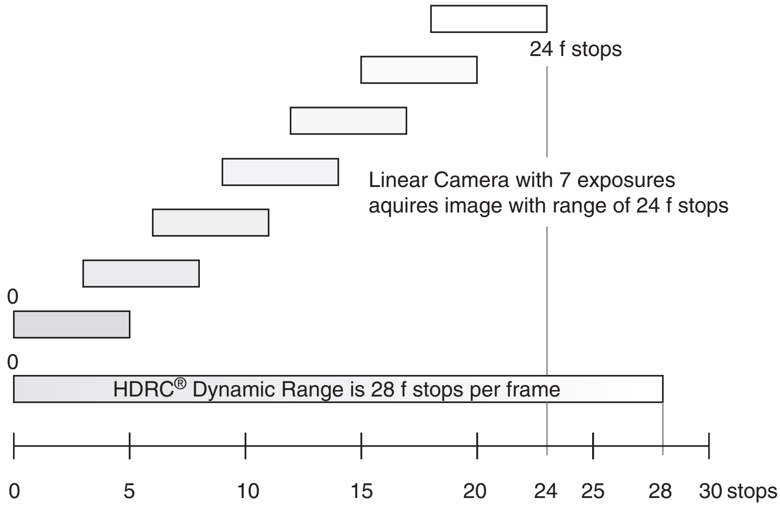

Fig. 1.5. HDR acquisition

with a conventional camera are staggered to obtain an HDR image file with a dynamic range of 23 f-stops, considered to be sufficient for a high-quality image file. Special stitching algorithms are required to fit the 7 files together into one comprehensive file. This art is covered extensively in HDRI Chap. 4 [3]. By comparison, HDRC VGAX sensors have been reported to have an instantaneous range of 8 orders of magnitude or 28 f-stops [2] so that one exposure captures all the information photometrically correct.

The optoelectronic conversion function (OECF) of the HDRC sensor, Fig. 1.1 is not only continuous and monotonic over so many orders of magnitude, it is moreover strictly logarithmic over the largest part. With its extremely high dynamic range, it reduces radically the design and operating problems of cameras:

> With HDRC, there is no need for aperture and/or exposure-(integration-) time control.

In other words:

With HDRC, no pixel and no frame can be over- or under-exposed.

In fact, as an accurate photometric sensor, it surpasses the human eye, which still needs pupil control covering about four f-stops to control the range of input intensities.

In the evolution of its senses, the human system has optimised its functionalities for efficiency and sensitivity among other features. The logarithmic conversion function is a wonderful case in point. In Fig. 1.1, we have plotted the OECF of the eye and the HDRC sensor on a comparable scale. Regarding the HDRC sensor, we see that the OECF maps an input range of 8 decades or 28 Bit to an output range of 10 Bit. So the first result of log conversion is information compression from 28 to 10 Bit:

The log converting HDR pixel provides natural bit-rate compression.

We may be concerned that this means a possible loss of information. As we shall see shortly, the opposite is true. We gain contrast sensitivity in darker, shaded parts of a scene. The so-called contrast sensitivity answers the most critical question:

What is the minimum change of the input grey value, which causes a just noticeable difference (JND) on the output?

The contrast sensitivity tells us, by how many percent we have to change an input value to obtain this JND on the output.

The human eye can distinguish changes in the input luminance, which are a little smaller than 1%, at sufficient light levels. In this region of "photopic" vision, this contrast sensitivity is independent of the input (grey) level as shown in Fig. 1.6. This is the region of the logarithmic response:

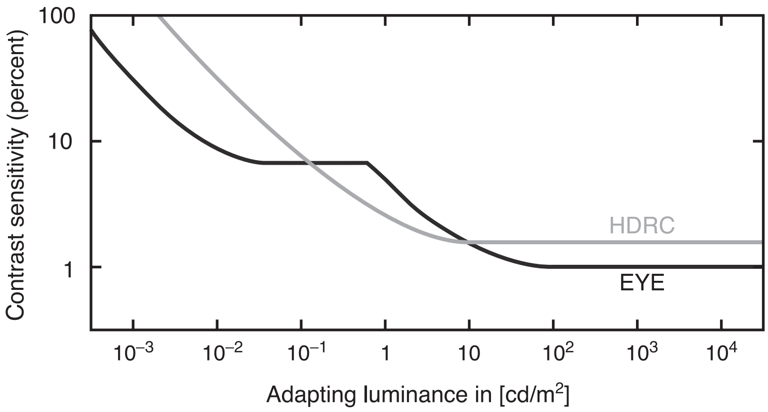

Fig. 1.6. The contrast sensitivity function (CSF) of the human eye and of an HDRC sensor with an input range of 7 orders of magnitude and a 10-Bit output

A logarithmic OECF offers constant contrast sensitivity independent of the input grey value.

It was Weber's finding in the 19th century that a natural response function has this unique feature that a constant "relative" change of the stimulus causes the just noticeable difference in the response. At lower light levels, the contrast sensitivity (CS) of the eye decreases: A relatively larger input change is needed to cause a JND on the output so that the spontaneous sensitivity of the eye is limited. However, with long-time adaptation, the sensitivity range then extends to very low light levels, as shown by the low-light tail of the CSF. We also find in Fig. 1.6 the CSF of the HDRC sensor. In the case of a 10-Bit output covering 7 orders of magnitude on the input, it is 1.5% and constant over 5 orders of magnitude. For the sensor with digital output, the JND is just 1 least-significant bit (LSB), often also called 1 digital number (DN) in the text. The sensitivity decreases at low light levels similar to the eye. However, the sensitivity is better than the spontaneous sensitivity of the eye. In the present discussion, we call sensitivity, in the language of film and standards, the minimum detectable signal. The relatively high performance of the electronic sensors is due to the properties and extensive technology development of silicon (Si) sensors. We put this in perspective relative to ASA and DIN sensitivity standards. For this purpose, we use the relation that a "green" light energy of 8 mlx s produces a just noticeable exposure (or a signal-to-noise ratio of unity) on a 100 ASA film.

The electronic receptor of choice, namely the Si photodiode, is limited in its detection capability by the so-called dark current. Its dominant component in miniature photodiodes is that along the perimeter of the photodiode. To get a feel for its significance, we assume that 50 dark electrons per micron are generated in 25 ms (a frame rate of 40 frames/s^{-1}) at a temperature of 50°C. The statistical variance of this number of electrons, the so-called shot noise, is given by the square root of this number (see Chap. 3). With a green illuminance of 20 mlx we would collect typically .1 electron/μm^{-2} in this frame time. The minimum detectable luminance is reached, when the number of photoelectrons has reached the number of electrons due to the dark shot noise. This would be the absolute physical limit. Other noise sources in an actual Si sensor make the total noise bigger and the sensitivity worse than the ideal presented here in Table 1.2.

This exercise shows that Si sensor technology has the potential for very good ASA/DIN ratings and that there are serious trade-offs between photodiode/pixel size and obtainable sensitivity. Progress in HDR capture in the future depends mostly on progress in low-light sensitivity. This exercise gives us a glimpse at where to look for improvements.

The extended range of high contrast sensitivity down to low light levels is essential for capturing details in the shady parts of HDR scenes. An example is shown in Fig. 1.7.

Table 1.2. The sensitivity of miniature Si photodiodes for different square sizes in terms of ASA and DIN ratings (see text for model assumptions)

Minimum detectable (mlx s): Unity signal-to-noise ratio (shot-noise limit)							
ASA	6,400	3,200	1,600	800	400	200	100
DIN	39	36	33	30	27	24	21
mlx s	0.12	0.24	0.5	1.0	2.0	4.0	8.0
Signal e μm^{-2}	0.25	0.5	1.0	2.0	4.0	8.0	16.0
Photodiode size (μm)	15.0	9.3	5.9	3.7	2.3	1.8	0.9
				Reference: Quantum efficiency: 50% Temperature: 50°C Dark electrons: 50e μm^{-1} Int. time: 25 ms			

Fig. 1.7. Portrait illuminated by intense side lighting, captured by an HDRC camera 512 × 256 pixels, 1/60 s, 1996

Inspite of the intense side lighting, the shaded half of the face shows the same amount of detail as the illuminated half, allowing safe contour extraction and face recognition.

We use this image to introduce another powerful feature eye-like HDR log capture:

Separation (or Disregard) of the Illuminant

We perceive real-world scenes predominantly by content (objects, contours, texture, color, motion) and less by actual and changing illumination. This has led to the feature of "Disregarding the Illuminant" in the context of describing the (HVS). In this introduction, we show practical examples where we correct or remove the incidental illumination in images recorded with HDR log-response sensors in real time to obtain image representations either pleasing for our eyes or suitable for robust machine vision.

Figure 1.8 shows four images of the Macbeth Chart recorded with an HDRC camera where the lens aperture was varied over 5 f-stops. Raw images on the left. On the right side, all pixels on the seemingly "over" or "under"-exposed frames have received an offset corresponding to the log of the numbers

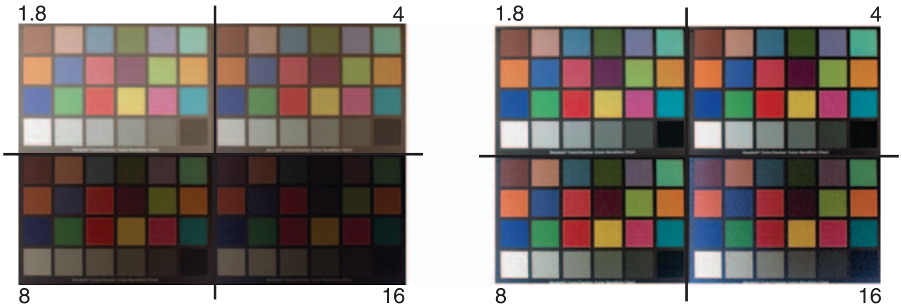

Fig. 1.8. Macbeth chart recorded with an HDRC camera with four apertures. Left: sensor output. Right: Pixel data shifted by common offsets

of f-stops necessary to get what we perceive as the "standard" representation of the chart.

A global "offset" correction, easily done in real time, restores the raw images on the left to the desired images on the right without any further processing or correction.

Leaving details to later chapters, let us consider here that the intensity incident on each pixel of our sensor is the product of the illumination and of the reflectance of that spot in our scene, which is imaged onto this pixel. Because of the logarithmic response of the HDRC pixel, the pixel output is the sum of log-luminance and log-reflectance. Therefore, an offset correction by log-luminance has the desired effect.

The result of this straightforward procedure is shown in Fig. 1.9 for the case of a rather spotty illumination of an object. When the profile of that illumination at the distance of that object, e.g. in front of an ATM, is known for that camera operation, the application of that profile as an offset produces an output image, which comes close to perfect studio lighting and which is ideal for feature extraction and recognition.

Fig. 1.9. One frame from a video sequence of a moving model with spot illumination recorded with an HDRC camera. On the left: The raw frame. In the middle: The log profile of the illumination recorded and stored after the installation of the camera. On the right: The illumination profile added as an offset in real time to the raw frame. Frame rate: 30 frames s^{-1}

We have shown in Figs. 1.8 and 1.9 that with HDRC eye-like video capture we can

Manipulate, Correct or Eliminate the Illuminant.

We just notice here that with the inverse process we can also custom-illuminate HDR log images.

What has been subtly introduced with these HDR color figures is another important feature of eye-like HDR acquisition:

Color Constancy

This phenomenon is the very complex capability of the HVS to perceive a color unchanged whether intensely or weakly illuminated and even if the spectrum of the illuminant changes. The latter is beyond the capability of present HDR capture. However, the first, namely rendering a color of any local area in a scene constant irrespective of whether it is brightly lit or shaded is evidenced by our two figures.

No local or global color correction of HDR log images is needed to compensate the illumination.

The offset by log luminance is the only operation performed to obtain the improved images in Figs. 1.8 and 1.9 Chapter 4 addresses HDR eye-like contrast and color in more detail.

Our paradigm of:

"HDR Photometric Video Capture mimicking the Eye's Response" offers unique features summarised here with references to the further treatment and important applications in this book:

1. Each pixel is a photometric log-compressing sensor with a dynamic range of 8 orders of magnitude (Chap. 2).
2. All pixels in the sensor have the same OECF (Chap. 2) due to calibration in real time.
3. Because of the very high dynamic range, there is no aperture or shutter-(integration) Time control (Chaps. 2, 3, 5, 6, 9 and 10)
4. The sensors have good low-light sensitivity, because the pixels operate basically at the shot-noise limit of their photodiodes (Chap. 3).
5. Constant contrast sensitivity over many decades of grey facilitates robust object detection in high-speed machine vision (Chaps. 5 and 6), in traffic (Chap. 8), in endoscopy (Chap. 9).
6. Pixel output is the sum of log-luminance and log-reflectance (R, G, B) (Chap. 4).
7. Log-luminance is available directly for each pixel (Chaps. 4, 9, 11 and 12).
8. Color per pixel is constant irrespective of luminance level (Chaps. 4, 7, 8, 9, 11 and 12).

The disruptive technology of HDR photometric video capture summarised in this chapter impacts the design of cameras and optics, and it enables entirely new, high-performance vision systems. HDRC video cameras are presented in Chaps. 5, 7, 8 and 9. Lenses for HDR video are discussed in Chap. 6. Some of the most demanding applications of HDR vision were selected for this book.The video-based guidance of aircraft is presented in Chaps. 5 and 6. Chapter 7 is an in-depth treatment of high-speed HDR machine vision for the inspection of complex, highly reflective parts. Chapter 8 offers a perspective on automotive driver assistance and the importance of HDR vision in this context. All the unique features of our new paradigm come to bear in new miniature cameras for minimally invasive endoscopy. It is a very special phenomenon that we return to the impaired eye what we have learned from the healthy eye, namely a subretinal implant, in which HDR log-compressing pixels generate a compatible stimulus for the retinal network as a partial replacement for deceased photoreceptors (Chap. 10).

Direct HDR photometric, log-compressing video capture with its film-like response, log luminance, natural bit-rate compression, constant contrast sensitivity and constant color has formidable implications on the video chain (Fig. 1.10) consisting of coding, transmission, tone mapping and display.

Chapters 11 (tone mapping) and 12 (coding) are critical reviews of these dynamic areas with an emphasis on comparative evaluations and original contributions, which have special affinity to HDR photometric capture. The HDR display presented in Chap. 14 is not only the one with widest dynamic range, but it is also the most illustrative inverse of the separation of log luminance and color in HDR video capture: What is displayed to us is basically the multiplication of regional luminance and high-resolution color.

To conclude this introductory chapter and in view of the many disciplines coming together in this book, we point the reader to the glossary of terms and abbreviations as well as to the comprehensive index in the back.

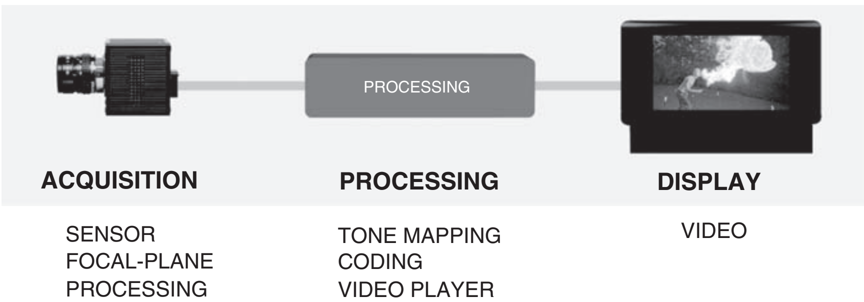

Fig. 1.10. The HDR vision chain from video capture through processing to display

References

1. Seger U, Graf HG, Landgraf ME (1993) Vision assistance in scenes with extreme contrast. IEEE Micro 13: 50–56
2. Seger U, Apel U, Höfflinger B (1999) HDRC imagers for natural visual perception. In: Jähne B, Haußecker H, Geißler P (eds) Handbook of computer vision and applications, volume 1 Sensors and applications. Academic, San Diego, pp. 223–235
3. Reinhard E, Ward G, Pattanaik S, Debevec P (2006) High dynamic range imaging. Elsevier, Amsterdam