Colorimetric and Resolution requirements of cameras

Alan Roberts

Addendum 32 : Assessment of a RED ONE camera

This document is a report of the results of tests that are the precursor of those described in the EBU technical document Tech3335. It is not an endorsement of the product.

Data for this addendum is taken from short examinations of a RED ONE camera, (build 14, February 14 2008, and build 17 on October 22, 2008).

The camera is of unusual form, breaking many of the "rules" for video or film cameras. It has a single large cmos sensor (super 35mm film, 24mm x13.5mm, with pixels at 5 micron spacing), 4520x2540 with Bayer pattern photo-sites. The normal lens mount is PL, taking 35mm movie-style lenses; a replacement lens mount is available enabling B4 (2"/3 video) lenses to be used, but that permits illumination of only the central part of the sensor (i.e. there is no lens in the adaptor). The camera system is intended primarily for high-end cinema production rather than television. At 4k resolution, it will make pictures at up to 60fps, while at 2k it will (currently) go up to 75fps. Later versions ay well go to higher frame rates.

The camera has a video viewfinder and HDSDI outputs. During recording and live viewing, this is only at 1280x720, but it replays recordings at 1920x1080. The camera's main output is on solid-state cards (only one slot in the camera), with considerable picture processing done in custom software (Redcine). The HDSDI feed is not of broadcast quality, so the only useable output from the camera is on the solid-state memory cards.

Since the camera records in a proprietary RAW format, with JPEG2000 compression, there are very few controls in the camera that affect the image. Most of the image control is done in post-production. Therefore, there was no attempt made to establish good settings for it, only to establish what it can do.

Colorimetric and Resolution requirements of cameras

Alan Roberts

Addendum 32 : Assessment of a RED ONE camera

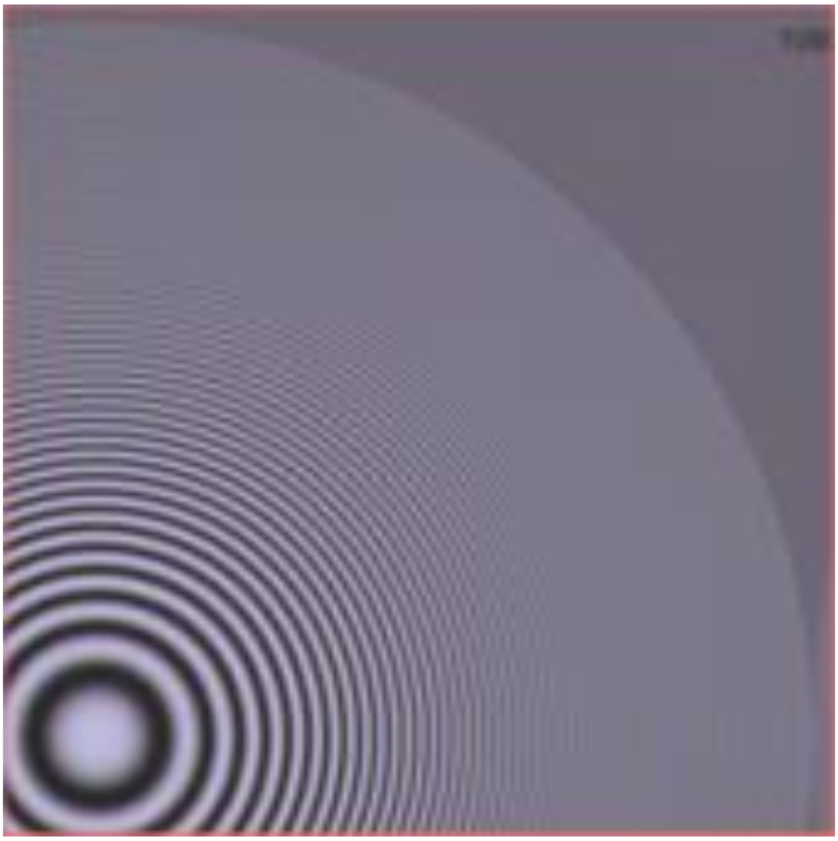
1. Resolution and Aliasing

The camera delivers images at 4k (4096x2304), 4kHD (3840x2160), 3k (3072x1728) and 2k (2048x1152) resolutions. None of these formats uses the full pixel-count of the sensor, some rows and columns, outside the nominal area, provide an over-scanned image so that the user can see scenic items coming into frame, just as in a film camera. Measurements were made using a circular zone-plate test chart, designed for 1920x1080 television. Although not ideal for testing the 4k resolution of the camera, the chart does contain small patterns which reach double HDTV resolutions, 3840x2160. Since BBC use of the camera is expected to be restricted to output at HDTV (1920x1080), the ways of achieving that performance were tested on this occasion.

1.1 Resolution at 4k (4096x2304)

The test chart was framed to fill half of the width and height of the 4k raster. Thus, the chart frequencies of 1920 pixels/width are actually about 3900 pixels/width in the image, approximately matching the pixel count of the image. In theory, this pattern should be fully resolved without aliasing, but in a Bayer-patterned single-sensor camera, the maximum resolution that can be expected to be delivered is about 70% is the pixel count, both horizontally and vertically, since the coloured pattern of pixels has to be decoded in order to generate the R, G and B signals.

This small part of the chart is one quadrant of a pattern, reaching 1920 horizontally and 1920 vertically, low frequency being in the at the large spot, and frequencies increasing linearly from it

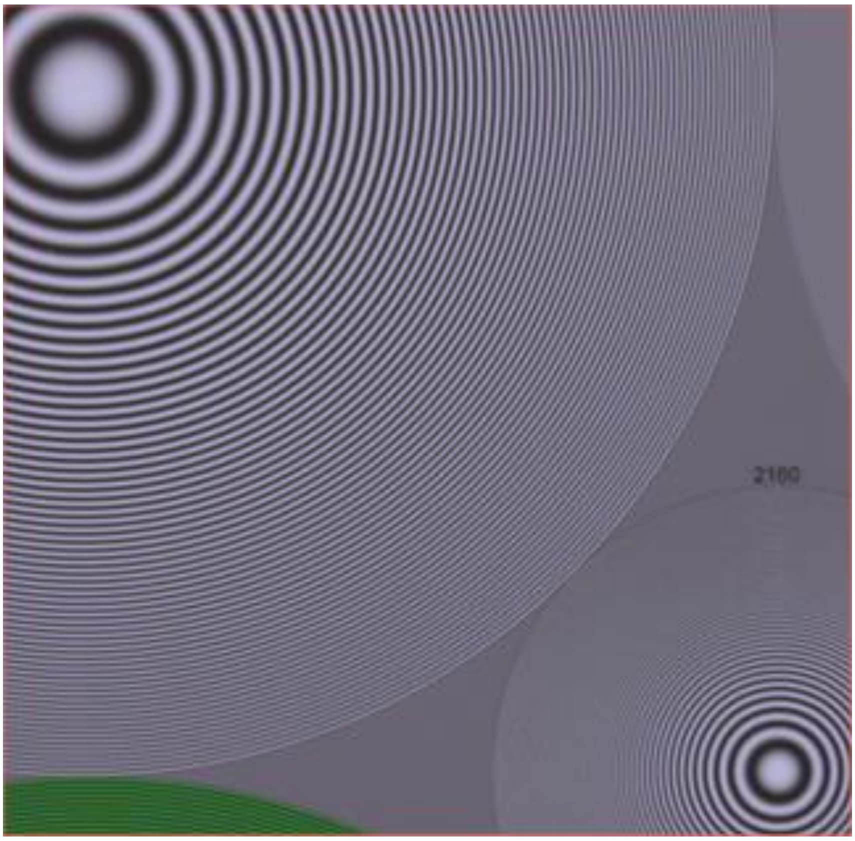


Resolution is clean up to about 67% of the camera's nominal limits of 4096x2304. This is perfectly normal for Bayer pattern decoding. Some coloured aliasing is visible above these frequencies along the horizontal and vertical axes, where the decoding is confused by the high spatial frequencies and produces low-frequency colours as well as high-frequency grey-scale. The presence of these aliases indicate that the camera's optical bi-refringent filtering is aimed at suppressing frequencies near and above the 4k limit, and not intended to eliminate frequencies above 2/3 of that count. This is both encouraging and disappointing, since it implies that the intention is to extract more high-frequency content than is being done at present, and that the current state of Bayer-pattern decoding is not good enough. But this is perfectly normal for any of the Bayer-patterned high-performance cameras.

It is difficult to see how this performance can be improved except by using far more complex decoding, probably with an element of adaptive filtering. Perhaps this will happen in Build 15, or in the final product.

1.2 Resolution at 4k, down-converted to 1920x1080

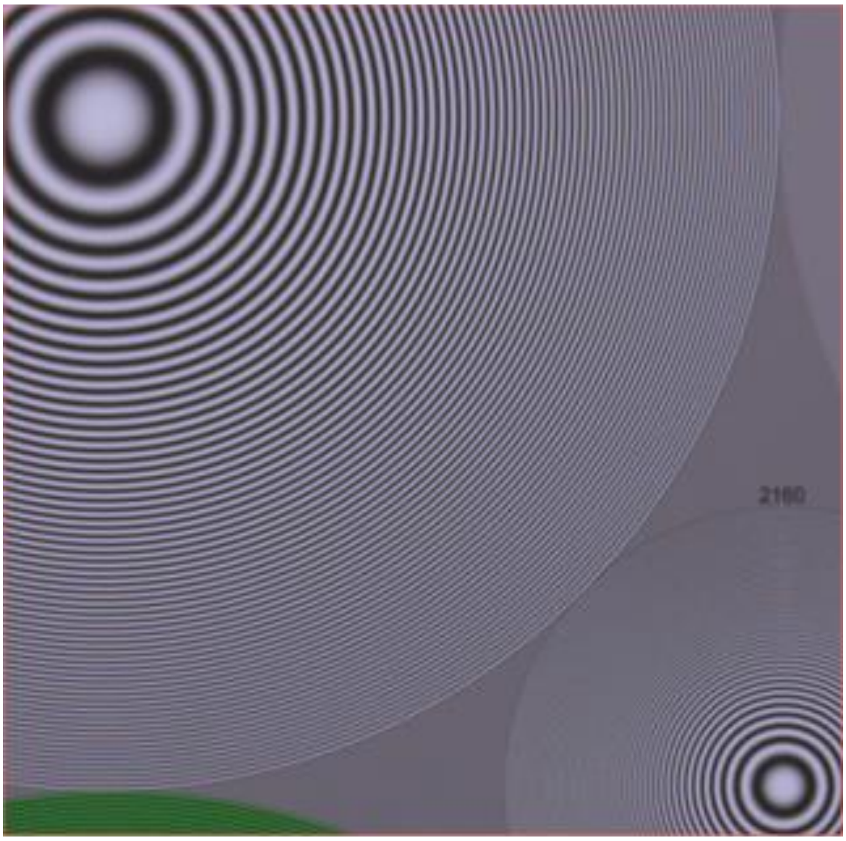
For this test, the chart was made to fill the 4k image, and the resulting capture was down-converted in Redcine, to 1920x1080 using the Lanczos algorithm, which was found to be the best on offer. Since the chart reaches only half the nominal frequency limit of the camera, the whole luma zone plate pattern should be resolved cleanly.



The chart contains two smaller patterns, reaching exactly double HDTV frequencies, 3840x2160; this shows the performance capturing the range of frequencies, and what happens to the higher frequencies which the camera can capture but are unwanted in the HDTV signal.

Clearly, the larger, 1920x1080 pattern shows clean resolution all over, but with contrast falling significantly but smoothly to the frequency limits. The smaller, double-frequency, pattern shows some coloured aliases from the higher frequencies that the camera can handle. These aliases are "burnt in" to the camera at 4k resolution, and cannot be eliminated by the down-conversion process.

The cure for this is to soften the lens with a mild diffuser, when shooting for HDTV, perhaps a Black Promist of density yet to be established. Ideally, such a filter should double the size of the lens' disc of confusion from a little over 5µm diameter to somewhat larger than 10 µm. Alternatively, use a super 16mm film lens rather than a 35mm movie lens, when shooting for HDTV.

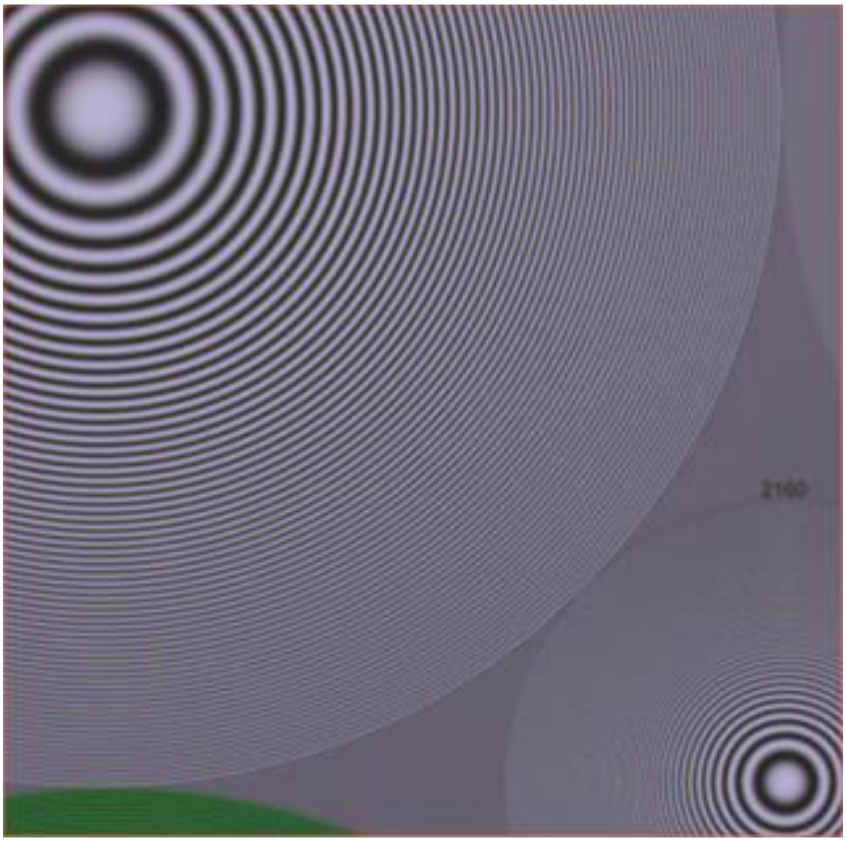


1.3 Resolution at 4kHD, down-converted to 1920x1080

Having set the camera to shoot at 4kHD (3840x2160), the chart was again exposed to fill the image, captured, and down-converted to 1920x1080.

The wanted resolution is slightly sharper, and coloured aliases are still visible in the smaller double-frequency pattern. This is hardly surprising, since there is a magnification change of only 0.9375 here, a drop of only 6.25%. Subjectively, there is no difference between this performance and that of shooting at 4k.

1.4 Resolution at 3k, down-converted to 1920x1080

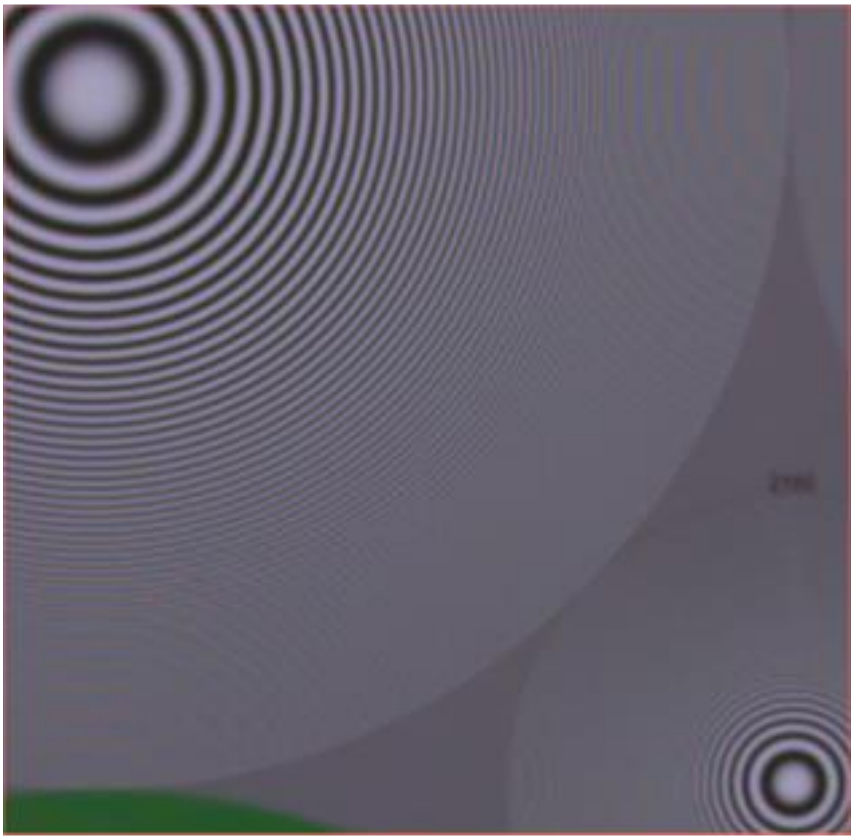


Repeating the process with the camera set to shoot at 3k (3072x1728), the chart was again exposed to fill the image, captured and down-converted to 1920x1080.

In this mode, the coloured aliasing in the smaller, double-frequency pattern have moved inwards, since the image is filling less of the camera's sensor and is presenting higher frequencies to it. Some coloured aliasing is now visible in the larger pattern.

Again, the cure would be to increase the size of the disc of confusion using either a light diffuser or a softer lens, such as a super 16mm format. However, aliasing at this level is not severe, and would cause no real problem in production but might be noticed during any compositing process or keying operation where accurate coloured edges need to be extracted.

1.5 Resolution at 2k, down-converted to 1920x1080



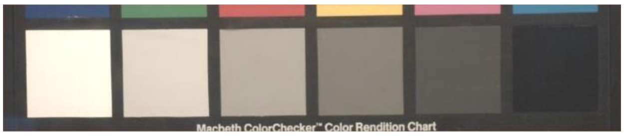
Once more, the camera was set to shoot at 2k (2048x1152) and the chart exposed to fill the frame.

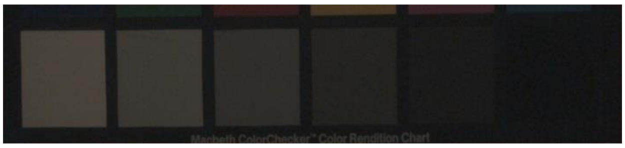
This time, the main pattern is soft, resolution at the edges of the main pattern is completely lost and there is significant coloured aliasing. Under these conditions, the camera is not really producing an HD image although it sits within an HDTV raster.

Since the chart is now almost exactly pixel-mapping it's resolution to the photo-site structure of the sensor, we are seeing the basic performance of the camera in 4k mode, effectively cropped down to 2k. The performance in this mode is little different from any consumer camcorder, with a 1920x1080 Bayer-patterned sensor, the coloured alias level is very similar. The result is not pleasing, and should not be considered suitable for use in HDTV production.

The down-conversion from 2k to 1920x1080 did not involve any scaling, the Redcine software simply cropped the image, which is probably the best thing to do.

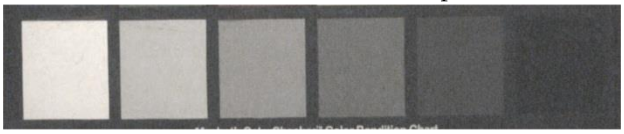
2 Contrast range



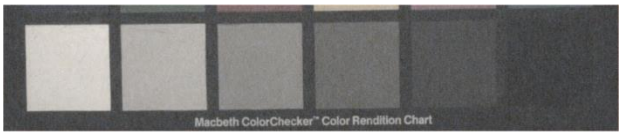


The camera was exposed to a Macbeth chart, at a range of exposures, and with the gain set to 1 (0dB). “Correct” exposure was judged to be T/5.6. Direct observation of the grabbed frames showed that exposures between T/3.5 (upper picture) and T/13 (lower picture) were at about the limits of the useable range. It is clear from this that the transfer characteristic used in the camera does not have any significant form of knee, compressing highlights into the top part of the coding range.

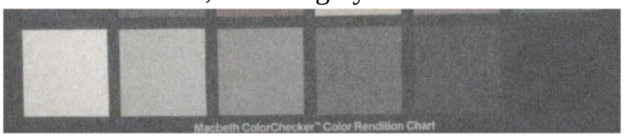
At T/3.5, the white patch is just clear of clipping, while at T/13 the black patch is just visibly different from the neighbouring grey patch. This is an exposure range of 4 photographic stops. The Macbeth chart has a range of reflectances over a range of 28.76:1, from white to black, which is a range of 4.86 stops. Thus the exposure range of the camera in this mode is 8.86 stops. This is, at first sight, a little worrying, since the best of the existing HDTV cameras are currently delivering contrast of 11 to 12 stops, with noise levels at about -55dB.



Further investigation shows, however, that the RED camera’s contrast range is actually rather better than 8.86 stops, since the perceptual limiting method (judging, by eye, when two adjacent patches on the grey scale are only just discriminable) normally used does not take into account the very low noise figure of the camera. Thus, in post-production, considerable gain can be used before noise becomes a problem. This illustration shows the T/3.5 frame treated to sufficient gain in post to make the grey scale “reasonably well exposed”, about 11dB gain. Clearly, this is still useable and noise is not a problem (since the luma noise has been worsened by 10dB and is now 51dB or so, around the performance level of existing HD cameras).



With this in mind, an exposure at T/16 was similarly treated to gain, to see what the noise looks like near black. Clearly, T/16 is still acceptable, the only difference being the rather higher noise level.



Similarly, exposure at T/22 is just about acceptable (gain about 14.5dB, so mid-grey noise level becomes about 45dB, still just acceptable), although the noise is now the major limitation. However, in a tight situation, such pictures could still be used for real production.

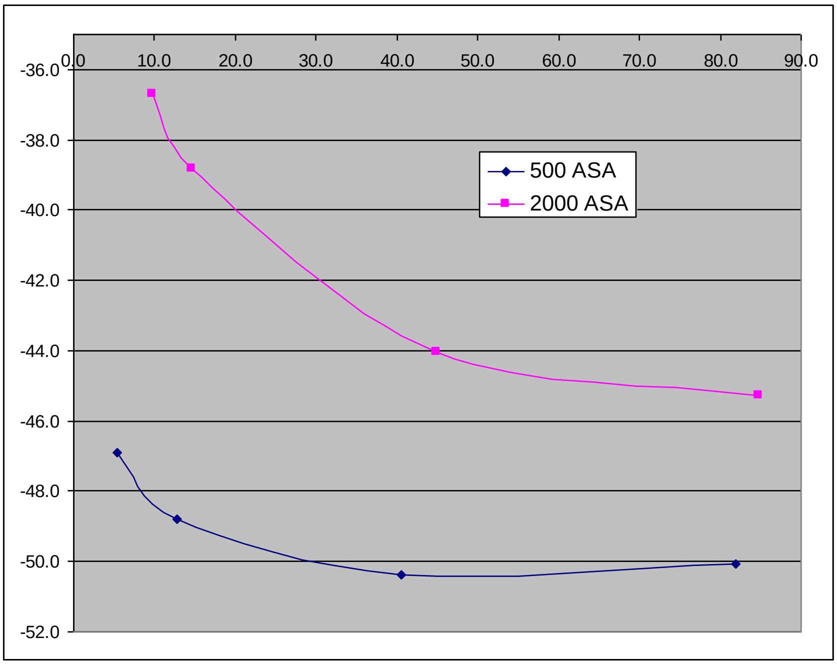
Thus, the exposure range is limited by the noise level, and not by the eye’s ability to discriminate adjacent shades. So, the exposure range can now be judged to be from T/3.5 to T/22, not T/13. This is a range of 5.5 stops, not 4. Add this to the 4.86 stops of contrast on the Macbeth chart and the usable range becomes 10.36 stops. Since the noise at T/22 is still not a major problem, a further half-stop is also probably acceptable, making a total range of about 11 to 12 stops. Coincidentally, this is the range usually claimed in all HD cameras.

3 Noise

Frames were captured at 4k resolution, and not down-converted. Four were captured with the camera gain set to unity (default) and the photographic speed set to 500ASA, and another four at 2000ASA. Captured frames were high-pass filtered and a further 6dB gain applied to help the calculations, the effect of this gain is eliminated in the results table.

Speed ASA	Video level %	Noise level, dB			Corrected for 6db gain,Y´
		R´	G´	B´	
500	5.4	-38.73	-42.01	-39.41	-46.89
500	12.8	-40.22	-44.05	-42.23	-48.80
500	40.6	-41.92	-45.52	-43.80	-50.37
500	82.0	-40.74	-45.79	-44.23	-50.05
2000	9.7	-28.29	-31.89	-32.85	-36.70
2000	14.7	-31.81	-33.41	-32.81	-38.83
2000	44.9	-37.22	-38.58	-36.13	-44.06
2000	84.8	-36.32	-40.76	-38.94	-45.27

The reason for capturing four frames at each setting was to establish how the noise changes with video level, since camera head noise is modified by the gain of the gamma correction curve.



The distribution of noise with signal level should follow the slope of the transfer curve, whether a gamma curve or logarithmic. Either way, there should be a difference of about 10dB between values near white and near 5~10%. The 2000ASA values clearly show this distribution, while the 500ASA values show much less correlation. This indicates that the internal data processing may well be providing the noise floor at a level near -60dB. One possible cause for this is a limitation of the bit-depth in the image processing, possibly at 14-bits in linear and/or 10-bits in log. This is one disadvantage of recording video using a log curve rather than a conventional gamma curve with knee function to extend the contrast range, since the gamma curve makes more efficient use of the available bit-depth, delivering up to 6dB better noise performance on test.

At lower ASA ratings (not specifically tested here), the noise levels should be significantly lower, e.g. 6dB better at 250ASA, about 56dB. Earlier, unpublished, tests showed a noise floor of about -60dB, which is typical for cameras in the top-end range of video capture. These tests confirm that view.

However, even at 500ASA the noise levels are acceptable, and do not compromise the contrast capture range, although 2000ASA may be excessive.

4 Fixed pattern noise

At 2000ASA, i.e. very high gain, and low exposure, some fixed pattern noise was just visible, as a series of vertical lines. On a grabbed still frame, the lines are almost impossible to see since they are at a lower level than the video noise. After a black-balance process, the visibility level dropped about 6dB. It is not clear what is causing this, and it is visible only in extreme conditions, so should not be a problem.

5 Rolling shutter

The sensor is not read out as a still frame, it is continuously scanned in a process known as “rolling shutter”. Technically, this is identical to the original scanning process used in tube-based cameras and displays. In the days when both camera and display were crt-based, there was no problem, since the scan position of the spot in the display was exactly matched to the scan position of the spot in the camera. However, when solid-state cameras emerged, the "rolling shutter" phenomenon was observed. When an object in the scene, or the whole scene, moved horizontally at speed, vertical lines appeared to lean backwards. This is because the camera was effectively taking still images, but the display was showing it continuously. Then, flat-panel displays came along, and all was well again, because both the camera and the display were working with what is effectively a series of still frames.

However, cameras are now being built, particularly those with cmos sensors, with continuous sensor scanning again. The result is that verticals in the image now lean into the motion when motion is rapid. The RED exhibits this effect.

Tests on moving material showed the effect, but not especially badly. In general, it is most visible when motion is very rapid (a horizontal translation rate of, say, 2 seconds per picture width). Normally, such motion would be blurred by camera integration, and probably by the defocusing effect of short depth of field. In normal use, it should not be a problem.

6 Conclusion

The RED camera is different from other cameras. It deals with images differently, and has to be treated differently to get the best from it. Its resolution and aliasing, when making 2k pictures directly, is rather disappointing, but at 4k it makes very nice pictures with content up to nearly 3k, while at 3k it is perfectly suited to making HDTV pictures. Best performance will be delivered if the lens has a resolution that is correct for the intended delivery standard; since the photo-sites are spaced at 5µm, the disc of confusion should be at least 7µm for 4k shooting, 10µm for 3k shooting, 15µm for 2k shooting. One way to achieve this is to use light diffusion on the lens, by an amount to be determined by experiment.

The noise and contrast range are, effectively, the same as the best of the current HDTV cameras, but contrast is handled rather differently, requiring some post-production effort to achieve the same contrast range. Nevertheless, the performance is as expected, given that all HD cameras use the same materials (silicon) in the sensors, and it the photo-chemistry of silicon that defines the limitations.

Although not assessed here in any detail, the HDSDI feed (1280x720 during recording, 1920x1080 during replay) is not fit for live recording for programme use, it contains too much aliasing. For normal production purposes, however, the viewfinder signal is perfectly adequate, but it should never be used as a programme feed.