Subpixel Image Scaling for Color Matrix Displays

Michiel A. Klompenhouwer*, Gerard de Haan†
Philips Research Laboratories, prof. Holstlaan 4 (WO-02), 5656 AA Eindhoven,the Netherlands,

Abstract

The perceived resolution of matrix displays increases when the relative position of the color subpixels is taken into account. 'Subpixel rendering' algorithms are being used to convert an input image to subpixel-corrected display images. This paper deals with the consequences of the subpixel structure, and the theoretical background of the resolution gain. We will show that this theory allows a low-cost implementation in an image scaler. This leads to high flexibility, allowing different subpixel arrangements and a simple control over the trade-off between perceived resolution and color errors.

1 Introduction

To generate full color images, matrix displays like Plasma Display Panels (PDPs) and Liquid Crystal Displays (LCDs) use three spatially displaced primary color subpixels (red, green and blue, *RGB*) per full color pixel. These subpixels are arranged in some repeating pattern, like the 'vertical stripe' arrangement, shown in Figure 1.

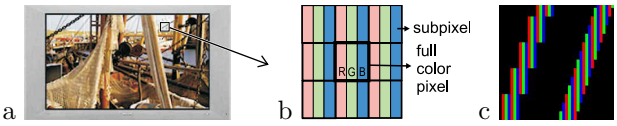


Figure 1: In a matrix display (a), each full color pixel consists of several primary color subpixels (b), here in the 'vertical stripe' arrangement. The question arises whether these subpixels can give extra resolution when the grouping into full color pixels is released, as illustrated by this simple example (c).

Each subpixel is given an intensity corresponding to the value of the color component at the corresponding full color pixel location in the image. When the subpixels are small enough, they are not individually visible at a normal viewing distance, so that the viewer will only perceive the resulting color (tristimulus value) for that location in the image: so-called 'color blending' occurs.

Figure 2 shows the basic diagram of the signal flow in a matrix display, as it will be discussed in this paper. In every displayed frame, each full color pixel needs values for all color components in the pixel. Therefore, the input signal must be processed and sampled such that there is a one-to-one correspondence between an input sample and a pixel of the display. In the addressing process, each input sample (*RGB* triplet for a specific location in the image) is directed to a particular full color pixel.

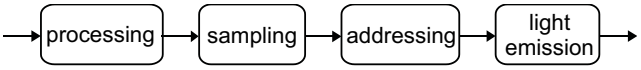


Figure 2: Basic diagram of signal flow in a matrix display system

This paper aims to determine a value for each subpixel such that the quality of the displayed image is maximized. In other words, the input image is reconstructed in the best possible way. Even if there is no one-to-one correspondence between the samples in the input image and the (sub)pixels of the matrix display, which is a common situation.

The subpixels represent a higher spatial resolution (three times in horizontal direction for the arrangement shown in Figure 1) than the (full color) pixel resolution [1, 2, 3, 4, 8]. However, we cannot neglect the color of the subpixels. Simply addressing the display with a monochrome image of a higher resolution would result in serious color artifacts, as will be shown in Section 3.1.

In references [2, 3, 8], it is suggested to take a weighted average of pixels in the triple-resolution input image. This will prevent color errors by spreading out the intensity of each input pixel equally over red, green

*michiel.klompenhouwer@philips.com
†g.de.haan@philips.com

and blue subpixels, creating 'virtual' or 'logical'[1] pixels. These *subpixel rendering* methods profit from an apparent resolution increase.

Given these subpixel rendering algorithms, the interesting question arises how much extra resolution is actually gained. This typically involves psychophysical modelling of human vision [5], or subjective testing [4, 12, 15]. Since this paper mainly deals with the consequences of the subpixel structure from a signal processing point of view, a detailed treatment of this question is considered outside the scope of this paper.

In the remainder of this paper, we will first analyse, in the frequency domain, the effect of the subpixel structure in displays. This analysis shows that we the increase in perceived resolution can be implemented at low additional cost in a flexible image scaler, allowing a straightforward control over the trade-off between color errors and sharpness. Finally, we will show results, on natural images and on text, and draw conclusions.

2 The resolution of a color matrix display

Let us calculate the frequency spectrum to analyse an image that is displayed on a matrix display with a vertical stripe subpixel arrangment[2]. Clearly, a matrix display can only approximate the light intensities corresponding to the original (space-continuous) image, since it generates a space-discrete signal.

Figure 3 shows the basic system model of the signal flow in a vertical stripe matrix display. Let us assume that the RGB signals are sampled at positions (x, y) that correspond to the center of a full color pixel at the display. These sampled signals, (r_s, g_s and b_s), form the input to the addressing process. This process translates into a spatial offset (delay) for each color. In this case $-\Delta x/3$ for R, and $+\Delta x/3$ for B, if Δx is the horizontal full color pixel distance. The translation of the signal amplitude to the physical light emission intensity, i.e. the reconstruction process, is carried out by the light emitting/transmitting area (aperture) of each (sub)pixel.

Formally, the sampling process results from multiplying the continuous signal with a 2-D series of δ-impulses, at intervals of Δx and Δy (vertically) [16]:

[1]The term 'logical' pixel has a different meaning in [3] from [14]. To avoid confusion, we will only use the terms '(full color) pixels' and 'subpixels'.

[2]We will use the example of the vertical stripe display, but the general approach is also applicable to other arrangements (see Section 4)

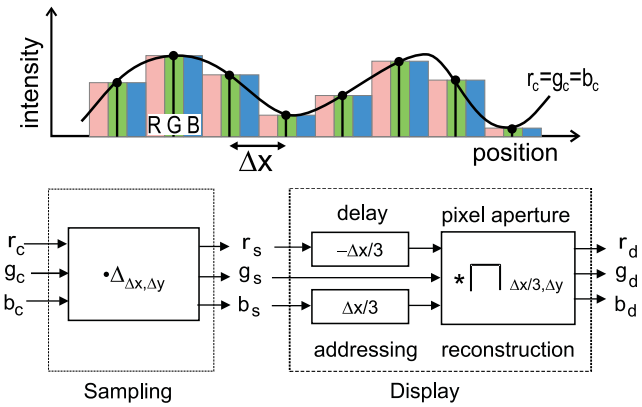


Figure 3: The RGB input (r_c,g_c and b_c) is sampled (to (r_s,g_s and b_s), and finally reconstructed by the display (r_d,g_d and b_d)

$$\Delta(x,y)_{\Delta x,\Delta y} = \sum_{k,l} \delta(x - k\Delta x, y - l\Delta y) \qquad (1)$$

The reconstruction process can be represented by a convolution of the sampled signal with an aperture function, $a(x, y)$, in this case a 2-D box function with width $\Delta x/3$ and height Δy:

$$a(x,y) = \sqcap(x,y)_{\Delta x/3,\Delta y} \qquad (2)$$

where

$$\sqcap(x,y)_{\Delta x,\Delta y} = \begin{cases} 1, & (-\Delta x/2 < x \leq \Delta x/2) \\ & \wedge\, (-\Delta y/2 < y \leq \Delta y/2) \\ 0, & \text{otherwise} \end{cases} \qquad (3)$$

We assume that the system is linear. Since video signals have a non-linear gamma pre-correction, and most displays have a non-linear electro-optical characteristic, this implies that proper gamma- and display correction are included in the system. For sake of simplicity of our analysis, we will not include these corrections.

In order to be able to indicate the perceptual impact of the subpixel structure and related algorithms, we transform the displayed image signal, r_d, g_d and b_d, to another color space. We choose the YUV space [11], because this has a separate luminance (achromatic) and chrominance channel. Similar results are obtained when any other color space is used that separates luminance and chrominance, or when the following matrix is adopted to accomodate other than the assumed primaries[3].

[3]This color transform is based on the color points of FCC (NTSC) primaries

$$\begin{bmatrix} y \\ u \\ v \end{bmatrix} = \begin{bmatrix} 0.30 & 0.59 & 0.11 \\ -0.17 & -0.33 & 0.50 \\ 0.50 & -0.42 & -0.08 \end{bmatrix} \begin{bmatrix} r \\ g \\ b \end{bmatrix} \quad (4)$$

We consider a monochrome image ($r_c = g_c = b_c = y_c \Leftrightarrow u_c = v_c = 0$), since luminance dominates the perceived resolution and it allows one to easily detect color errors.

We start with the continuous image: $r_c(x,y)$, $g_c(x,y)$, $b_c(x,y)$, and sample at the same position (see also Figure 3):

$$\begin{aligned} r_s(x,y) &= r_c(x,y) \cdot \Delta(x,y)_{\Delta x,\Delta y} \\ g_s(x,y) &= g_c(x,y) \cdot \Delta(x,y)_{\Delta x,\Delta y} \\ b_s(x,y) &= b_c(x,y) \cdot \Delta(x,y)_{\Delta x,\Delta y} \end{aligned} \quad (5)$$

Following our display model, the sampled RGB signals are now shifted relative to each other and convolved with the pixel aperture.

$$\begin{aligned} r_d(x,y) &= r_s(x+\Delta x/3, y) * a(x,y) \\ g_d(x,y) &= g_s(x,y) * a(x,y) \\ b_d(x,y) &= b_s(x-\Delta x/3, y) * a(x,y) \end{aligned} \quad (6)$$

Combining Eq. (5) and (6), and computing the Fourier transform ($\mathcal{F}\{\ldots\}$)

$$R_d(f_x,f_y) \quad = \quad \mathcal{F}\{r_d(x,y)\} \quad (7)$$

(similarly for G_d,B_d) results in the frequency spectrum of the displayed image:

$$\begin{aligned} R_d(f_x,f_y) &= \left[R_c * \Delta_{\frac{1}{\Delta x}\frac{1}{\Delta y}}\right] A(f_x,f_y) e^{2\pi i f_x \Delta x/3} \\ G_d(f_x,f_y) &= \left[G_c * \Delta_{\frac{1}{\Delta x}\frac{1}{\Delta y}}\right] A(f_x,f_y) \\ B_d(f_x,f_y) &= \left[B_c * \Delta_{\frac{1}{\Delta x}\frac{1}{\Delta y}}\right] A(f_x,f_y) e^{-2\pi i f_x \Delta x/3} \end{aligned} \quad (8)$$

where

$$R_c = R_c(f_x,f_y) = \mathcal{F}\{r_c(x,y)\} \quad (9)$$

is the Fourier transform of the continuous image (similiarly for G_c and B_c) and

$$A(f_x,f_y) = \mathcal{F}\{a(x,y)\} = \frac{sin(\pi f_x \Delta x/3)}{\pi f_x \Delta x/3} \frac{sin(\pi f_y \Delta y)}{\pi f_y \Delta y} \quad (10)$$

is the Fourier transform of the pixel aperture. Now we calculate Y_d:

$$\begin{aligned} &Y_d(f_x,f_y) = \\ &0.30R_d(f_x,f_y) + 0.59G_d(f_x,f_y) + 0.11B_d(f_x,f_y), \end{aligned} \quad (11)$$

which simplifies to (omitting (f_x,f_y) for compactness):

$$Y_d = [Y_c * \Delta_{f_{sx},f_{sy}}] A \Phi_Y \quad (12)$$

where $f_{sx} = \frac{1}{\Delta x}$ and $f_{sy} = \frac{1}{\Delta y}$ are the horizontal and vertical sampling frequencies, respectively, and Φ_Y is defined by:

$$\Phi_Y(f_x,f_y) = 0.30e^{2\pi i f_x \Delta x/3} + 0.59 + 0.11e^{-2\pi i f_x \Delta x/3} \quad (13)$$

Similarly, U_d and V_d result as (note that $U_c = 0$ and $V_c = 0$):

$$\begin{aligned} U_d &= [Y_c * \Delta_{f_{sx},f_{sy}}] A \Phi_U \\ V_d &= [Y_c * \Delta_{f_{sx},f_{sy}}] A \Phi_V \end{aligned} \quad (14)$$

with

$$\begin{aligned} \Phi_U(f_x,f_y) &= -0.17e^{2\pi i f_x \Delta x/3} - 0.33 + 0.5e^{-2\pi i f_x \Delta x/3} \\ \Phi_V(f_x,f_y) &= 0.5e^{2\pi i f_x \Delta x/3} - 0.42 - 0.08e^{-2\pi i f_x \Delta x/3} \end{aligned} \quad (15)$$

Figure 4 shows a plot of the amplitude of the horizontal spectrum of the displayed image using pixel sampling, according to Eq. (12) and (14). In order to show amplitude of baseband and repeats, the spectrum (Y_c) of the original image has been assumed flat inside, and zero outside the baseband[4] ($-0.5f_s < f_x < 0.5f_s$).

Remember that the displayed image is further 'filtered' by the human visual system, causing the highest frequencies to become invisible. The range of visible frequencies depends primarily on the relationship between viewing distance and pixel pitch (Δx). The relation between the display spectrum and its perception therefore also depends on the viewing distance and the pixel pitch [5]. However, the sample frequency f_s gives us an indication of the optimal viewing distance. This frequency, which corresponds to the frequency of the (sub)pixel structure itself, will ideally be (slightly) higher than the threshold of visibility, while keeping the Nyquist frequency ($f_s/2$), the highest frequency still carrying valid, non-aliased, image information, within the range of visibility. If this is true, the display will have, neither too little, nor too much resolution. In other words, the pixel structure becomes visible, or valid image information is not visible. The relationship between displayed images and the human visual system characteristics is well studied [19], also with respect to the perceived quality of color matrix displays [12, 15], and a detailed discussion is considered to be outside the scope of this paper.

Figure 4 shows a number of properties of the display. First, reconstruction of the original image is not perfect. The baseband is attenuated, and repeats are present. For example, the non-zero amplitude at the

[4] f_{sx} is abbreviated to f_s

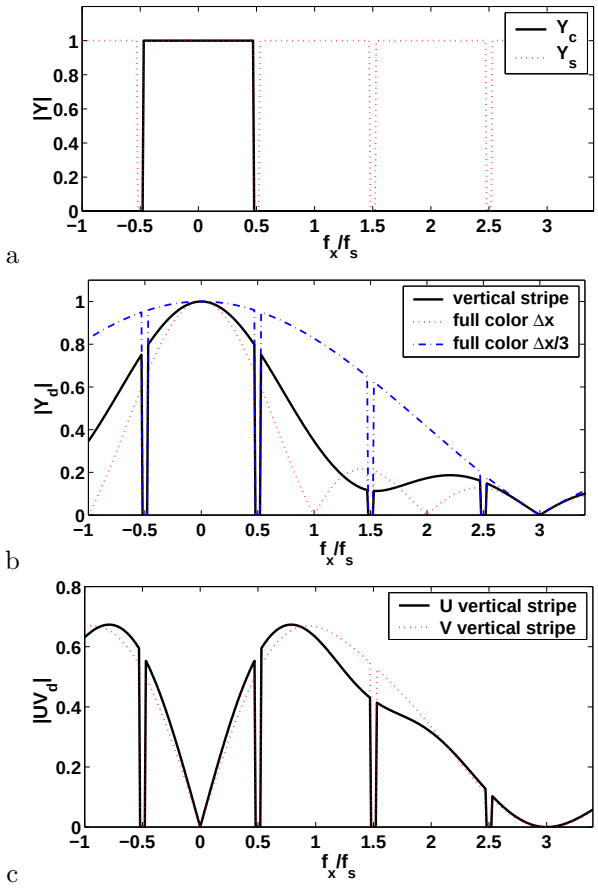


Figure 4: Horizontal frequency spectra of an image on a vertical stripe display. a) continuous (Y_c) and sampled (Y_s) images, b) luminance of displayed image (Y_d), c) chrominance of displayed image (UV_d). The non-zero frequencies in Y_d,U_d and V_d outside the baseband demonstrate an imperfect reconstruction (see text).

sampling frequency indicates that pixel structure is present even in uniform areas. Clearly, perfect reconstruction is not possible for matrix display systems, because this would require a *sinc*-shaped pixel aperture, which is not only infinitely wide, but also requires negative light intensities. This practical limitation is reflected in the so-called Kell factor, which will be further discussed in Section 3.2.

We further observe that the luminance spectrum of the pixel sampled image lies largely in between the spectra of displays with non-displaced subpixels[5] of widths Δx and $\Delta x/3$. Figure 5 shows why: The luminance profile of the combined *RGB*-aperture is somewhere in between these two examples. When *RGB* do not have equal value, the resulting profile will vary accordingly.

[5]This could e.g. be a projection or a color-sequential display

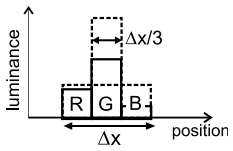


Figure 5: Luminance profile of a vertical stripe full color pixel (solid lines) and (dashed lines) pixels, with non-displaced subpixels, with widths Δx and $\Delta x/3$.

Figure 4 also shows the effects on the chrominance. Since we started with a monochrome image ($U_c = V_c = 0$), any non-zero U,V is an error. All frequencies except $f_x = 3nf_s$ ($n \in \mathbb{Z}$) have non-zero U,V. This is caused by the constant 'misconvergence' between *RGB* of 1/3 of a pixel, and is worst for high frequencies. Of course, for high frequencies, in particular the sampling frequency itself, we expect to see chrominance 'errors', because these frequencies correspond to the pattern of the highly saturated, individual subpixels. The human visual system is insensitive to these high chrominance frequencies, because we assumed that the subpixel colors are blended (see Section 1).

3 Subpixel sampling

Taking into account the position of the subpixels is equivalent to sampling the signal at the actual subpixel position of each color: subpixel sampling, as illustrated in Figure 6. This seems very straightforward, and indeed it has been proposed before [1, 7, 14]. As a matter of fact, color Cathode Ray Tubes (CRTs) apply this principle for decades, by virtue of the shadowmask [17, 18]. Yet for CRTs, the potential resolution gain is prevented in practice by the electron spot profile. The size of the electron spot, besides imposing limitations of the gun focus, is also limited by scan Moiré considerations. However, for matrix displays, these limitations do not apply, and the subpixel structure can be exploited to optimize the resolution of these displays.

The subpixel rendering methods in references [2, 3, 8] are also based on this principle, be it implicitly through the use of 'logical' pixels. Our analysis is different for two main reasons. First, we perform the general subpixel sampling on a continuous image, so we do not pose constraints on the input resolution. Second, we make an explicit split between the sampling process and the preceding filtering, i.e. we do not try to find 'logical pixels'. Finally, the step towards practical input images will become clear when we consider scaling in Section 5.

To calculate the resulting spectrum, we start again with

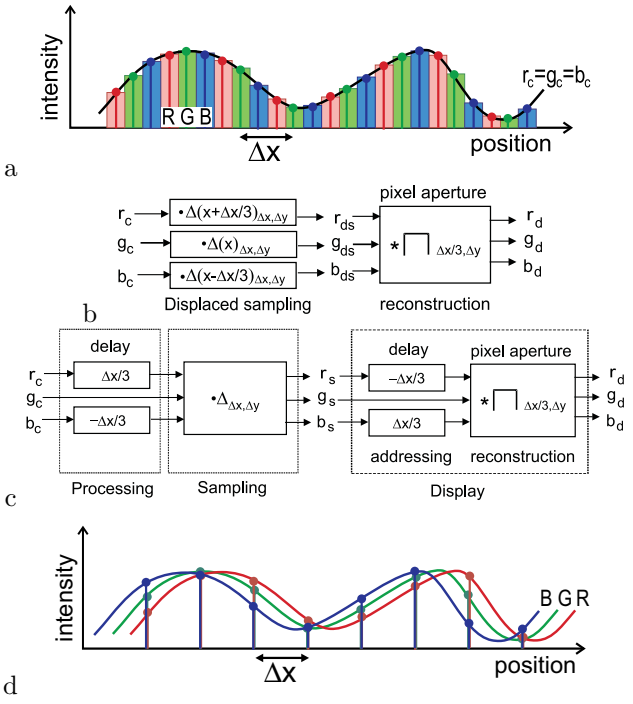


Figure 6: a) Subpixel sampling and reconstruction of the RGB input, b) system diagram using displaced sampling, c) equivalent system using RGB delay, d) sampling of delayed RGB signals

a continuous image signal r_c, g_c and b_c, which is now sampled at different positions for each color, according to Figure 6. This can be described by 'displaced sampling' [2]:

$$\begin{aligned} r_{ds}(x,y) &= r_c(x,y)\Delta_{\Delta x,\Delta y}(x+\Delta x/3,y) \\ g_{ds}(x,y) &= g_c(x,y)\Delta_{\Delta x,\Delta y}(x,y) \\ b_{ds}(x,y) &= b_c(x,y)\Delta_{\Delta x,\Delta y}(x-\Delta x/3,y) \end{aligned} \quad (16)$$

However, this also removes the addressing stage from our display model. An equivalent description is obtained by adding a delay before sampling:

$$\begin{aligned} r_s(x,y) &= r_c(x-\Delta x/3,y)\Delta_{\Delta x,\Delta y}(x,y) \\ g_s(x,y) &= g_c(x,y)\Delta_{\Delta x,\Delta y}(x,y) \\ b_s(x,y) &= b_c(x+\Delta x/3,y)\Delta_{\Delta x,\Delta y}(x,y) \end{aligned} \quad (17)$$

since

$$\begin{aligned} r_{ds}(x,y) &= r_s(x+\Delta x/3,y) \\ g_{ds}(x,y) &= g_s(x,y) \\ b_{ds}(x,y) &= b_s(x-\Delta x/3,y) \end{aligned} \quad (18)$$

This is also shown in Figure 6. Both discriptions result, after a similar calculation as in Section 2, in the YUV spectrum (Y_d, U_d, V_d) of the displayed image (again using $R_c = G_c = B_c = Y_c$):

$$\begin{aligned} Y_d &= [Y_c * (\Delta_{\frac{1}{\Delta x},\frac{1}{\Delta y}}\Phi_Y)]A \\ U_d &= [Y_c * (\Delta_{\frac{1}{\Delta x},\frac{1}{\Delta y}}\Phi_U)]A \\ V_d &= [Y_c * (\Delta_{\frac{1}{\Delta x},\frac{1}{\Delta y}}\Phi_V)]A \end{aligned} \quad (19)$$

Φ_Y,Φ_U and Φ_V are defined by Eq. (13) and (15). Comparing Eq. (19) and Eq. (12), we can see that the complex phase factor has moved inside the convolution, which implies that each repeat is attenuated as a whole by the values of this factor at the repeat position.

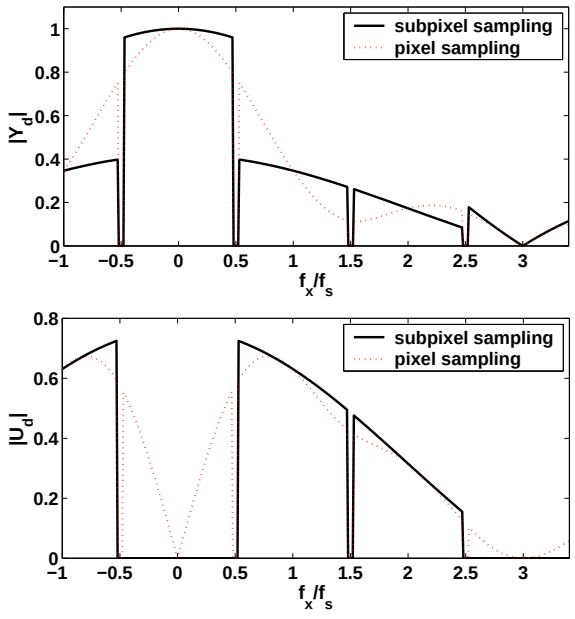


Figure 7: Horizontal frequency spectrum of an image on a vertical stripe display, using subpixel sampling. a) luminance, b) chrominance (U only)

Figure 7 shows the frequency spectrum of the displayed image using subpixel sampling, according to Eq. (19).

The correspondence between pixel and subpixel sampling is visible at multiples of the sample frequency ($f_x = nf_s, n \in \mathbb{Z}$). These frequencies correspond to an image with only a DC component, i.e. a uniform area, where it should not matter whether the RGB signals are sampled according to the subpixel positions or not (the pixel structure has not changed), and indeed the two spectra are identical there.

At other frequencies, the differences between the pixel- and subpixel sampled spectra are substantial: the subpixel sampled luminance spectrum is less attenuated in the baseband and the repeats are more suppressed. The baseband attenuation of the pixel sampled image is caused by the larger reconstruction 'unit' (see Figure

5). The lower repeat amplitude in the subpixel sampled luminance spectrum is caused by a shift of the repeats from the luminance to the chrominance signal. This also shifts aliasing from the luminance to the chrominance, resulting in false colors for image parts with frequencies outside the baseband, as shown in Figure 8. This figure shows the luminance, Y, and the chrominance (indicated for one, U, of the components only) of a zoneplate image, on a simulated vertical stripe display using both types of sampling. For example, a pure black and white signal at a frequency $f_x = f_s$, will appear as a constant color. This has been reported before [2, 9] as a potential problem, but it is also an improvement, as will be explained in the next section.

Figures 7 and 8 also show that subpixel sampling has actually *reduced* the chrominance errors inside the baseband. This is explained by the fact that sampling RGB at the subpixel positions, corrects for the 'misconvergence' that otherwise occurs.

3.1 Color aliasing

Due to the different sampling phases of RGB, the (lowest order) repeats in the frequency spectrum of the displayed image have partly[6] shifted from the luminance to the chrominance, which is clearly demonstrated by Figures 7 and 8. In [9] this 'discoloration' was shown, and it was also shown that the discoloration disappears when the image moves at a certain velocity. However, we like to prevent color artifacts also in image parts that move at other than this particular velocity. An appropriate filter prior to sampling, to suppress at least the most disturbing aliasing, is therefore an important feature in all subpixel rendering methods [2, 3, 8], as indeed it is in all sampling systems.

Clearly, subpixel sampling does not provide the threefold resolution we might hope for. The real improvement is found at frequencies much closer to the baseband.

The spatial frequency of the aliased color signal will increase as the signal frequencies come closer to the baseband. It is well known that the human visual system is less sensitive to high frequency chrominance errors than to high frequency luminance errors [5, 19], so that this high frequency chrominance aliasing can be tolerated to some extent.

Consequently, the cut-off frequency of the anti-alias filter can be extended beyond the Nyquist limit, which gives the highest aliasing frequency, but certainly not up to the sampling frequency, which causes a DC-alias frequency, i.e. a constant color error. Effectively a trade-off results between color errors, when too many high frequencies are passed, and unsharpness, when too many high frequencies are suppressed. A filter that suppresses frequencies close to the Nyquist limit can even render the subpixel sampling useless - apart from the misconverge correction - because the largest differences between the two sampling methods are found in this region.

Note that the Nyquist limit itself has not shifted: frequencies above the Nyquist limit still cause aliasing, although it has partly shifted from the luminance to the chrominance. However, this released constraint on the anti-alias filter around the Nyquist limit can not fully explain the apparent resolution increase associated with subpixel sampling. There is also a difference for frequencies below the Nyquist frequency, as we shall discuss in the next section.

3.2 Increased Kell factor

Due to the imperfect reconstruction, baseband frequency components ($f_x < f_s/2$) still have a version mirrored in the Nyquist frequency ($f'_x = f_s - f_x$). The interference of these two components results in a 'beat' pattern:

$$cos(f_s - f_x) + cos(f_x) = 2cos(f_s/2)cos(f_s/2 - f_x), \tag{20}$$

of which the frequency $(f_s/2 - f_x)$ decreases when the baseband component and its mirror approach the Nyquist frequency. This is illustrated in Figure 9 for a zoneplate image reconstructed with square pixels.

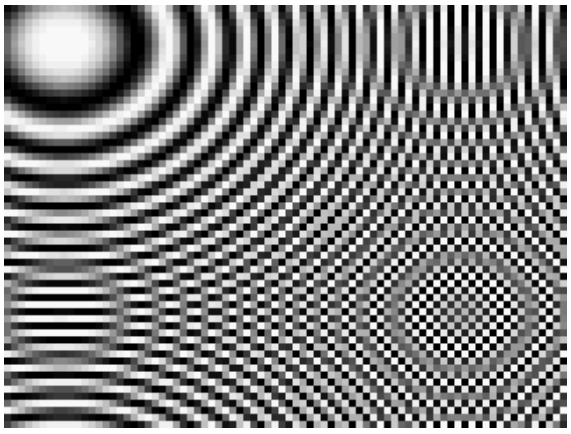


Figure 9: Zoneplate, with square pixels, showing a beat pattern around the Nyquist frequency (which corresponds to alternating black and white lines).

[6] Only if the luminance contributions of RGB would be equal, the shift would be complete: no luminance aliasing below spatial frequencies of $3f_s/2$.

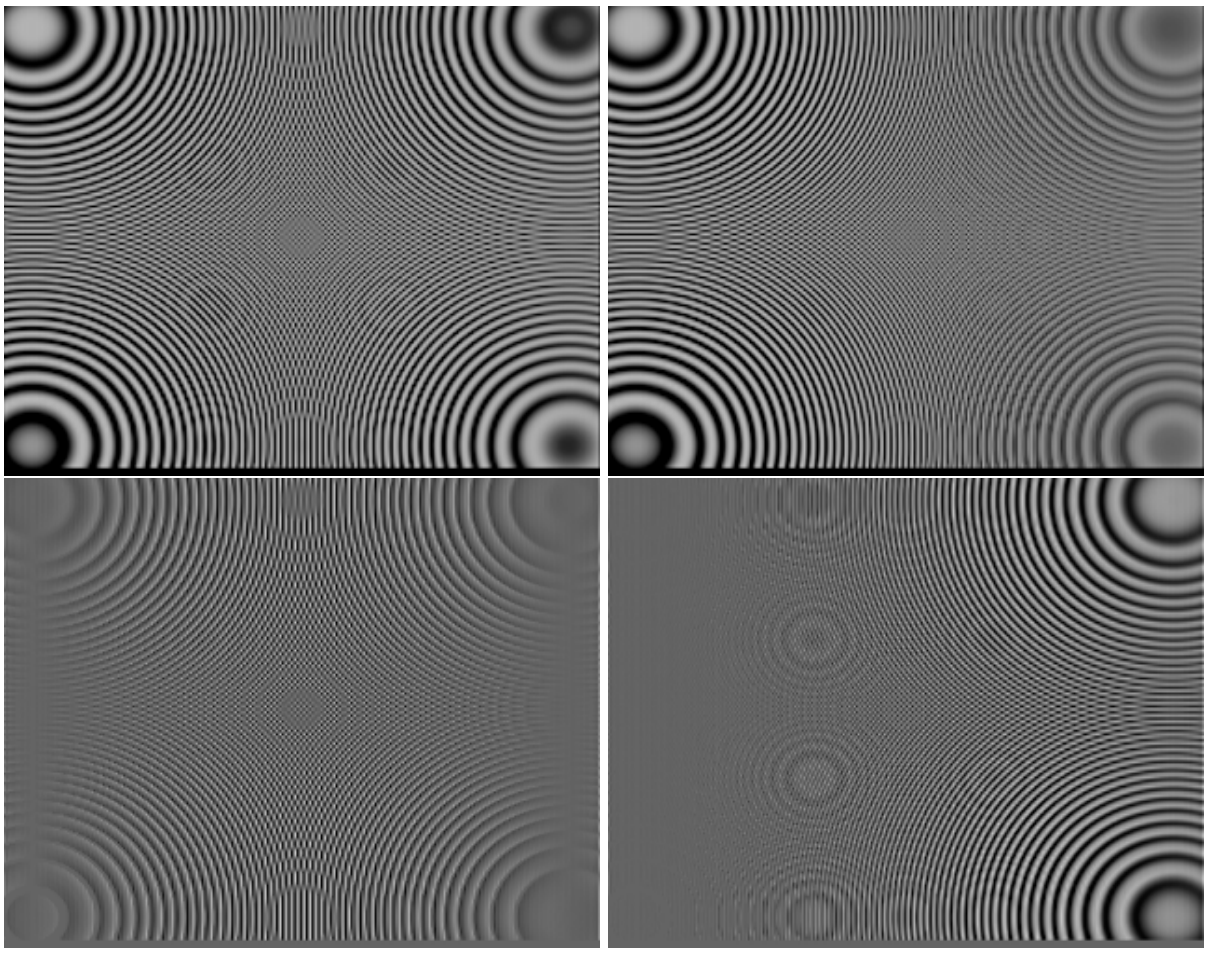


Figure 8: Zoneplate (radial frequency sweep, here from $f(x,y) = 0$ to $f(x,y) = f_s$) on a vertical stripe display. Left: pixel sampling, right: subpixel sampling, top: luminance (Y), bottom: chrominance (U)

The beat pattern reflects that the highest frequencies cannot be represented equally well in all phases. This is not a form of aliasing, but is entirely due to the imperfect reconstruction, i.e. the missing interpolation filter to reconstruct the continuous signal from the dicrete signal.

At some point, this beat pattern will dominate the perception, and reduce the practical use of passing this part of the baseband spectrum to the display. This is the origin of the Kell factor [10], which indicates the fraction of the baseband that is effectively available according to perception experiments. Although the Kell factor was introduced to balance the horizontal bandwidth in the Cathode Ray Tube (CRT) with the number of scanned lines (resulting in a typical Kell factor of 0.7), the imperfect reconstruction filtering also accounts for such a factor in matrix displays!

Figure 10 shows the upper part of the baseband from figure 8, and the beat patterns are also visible here. With subpixel sampling, we can see that, due to the increased amplitude difference between baseband and repeat, the amplitude of the beat frequencies will be reduced.

Stated differently, the beat frequencies will be shifted from the luminance to the chrominance, which reduces their visibility. This translates into an increased Kell factor, without changing the physical aspects of the display. Even though this does not increase the resolution in terms of number of pixels, it does increase the perceived sharpness. Therefore, subpixel sampling allows us to display the frequencies below the Nyquist limit of the display with less distortion.

This was also found in [4], where the 'pixel structure noise', related to the beat patterns, was measured by determining the maximum frequency that a viewers panel could recognize. The 'noise' was found to be less for grayscale images than for images of a single primary color.

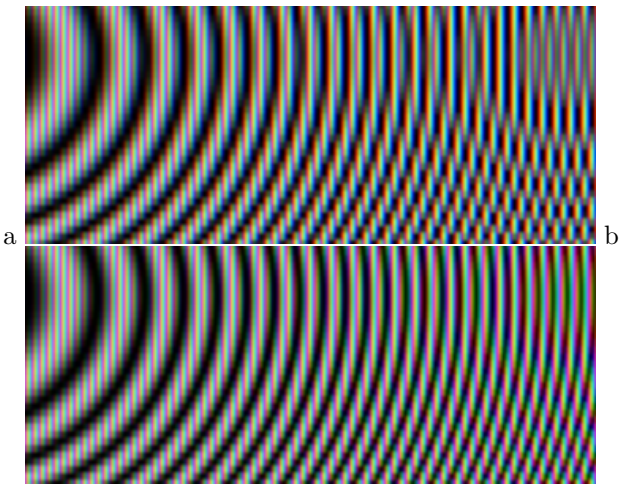


Figure 10: Zoneplate (part of baseband), a) pixel sampling, b) subpixel sampling

Comparing the diagonal high frequency parts of Figure 10 and Figure 1c, we can see that the beat frequencies and the jagged lines that are usually associated with a low resolution (or poorly anti-aliased) image, have the same origin.

4 Alternative subpixel arrangements

The general principle for alternative subpixel arrangements is the same. The lower repeat spectra are shifted from the luminance to the chrominance, giving less distortions inside, and reduced aliasing outside the baseband.

Consider for example the 'Delta-Nabla' subpixel arrangement, also called (Delta) Triad, shown in Figure 11. Here, the subpixels are arranged in a two-dimensional pattern. More precisely, the subpixels are still displaced in horizontal direction ($2/3\Delta x$), but the pattern is also displaced (Δx) from line ($\Delta y/2$) to line, to create a diagonal[7] sampling lattice for each color.

A formal description is more complicated and adds little to our understanding, so we directly show the result using the zoneplate image in Figure 12.

The right part of Figure 12 shows the subpixel sampled image. Compared to Figure 8, the repeats are shifted from the luminance to the chrominance, and furthermore, the repeats in the high vertical frequency part of the spectrum, have moved to a high diagonal frequency location. When there is no subpixel sampling, i.e. the 'Delta-Nabla' addressing scheme is used with *RGB* signals that are sampled at the same location, the Delta-Nabla arrangement does not increase the resolution, as shown in the left part of Figure 12. The repeats are at the same location as for the vertical stripe arrangement, although the visibility of the pixel structure has been reduced. There are actually more artifacts, such as jagged edges and discoloration around the Nyquist limit. To exploit the extra resolution that the Delta-Nabla arrangement offers, we have to use subpixel sampling.

[7] Or hexagonal, depending on the ratio $\Delta x/\Delta y$

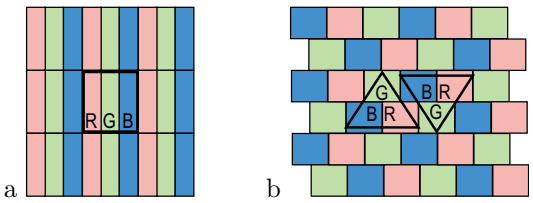


Figure 11: a) Vertical stripe and b) Delta-Nabla subpixel arrangements with the same full color pixel size

Also for other subpixel arrangements, notably the 'PenTile' arrangement [3], resolution benefits are only perceived if the sampling is adapted to the subpixel structure.

5 Subpixel image scaling

In the previous section, we explained some of the benefits of subpixel sampling. In this section, we will propose an efficient and flexible implementation.

For the spectrum analysis in Sections 2 and 3, we started with a continuous image, to generalize the subpixel sampling principle. In practical situations, we probably only have a digital signal, at a sampling grid that does not match our display. Therefore, the input image needs to be scaled to the display grid, which can then be used in the one-to-one mapping of the display addressing.

Image scaling, or sampling rate conversion, is a basic signal processing technique for multirate sytems [6, 13], and can be described, for rational scaling factors, as a cascade of upsampling (zero insertion), lowpass (anti-alias) filtering, and downsampling (decimation), as shown in Figure 13.

These operations are also required for subpixel sampling and a very straightforward adaptation of an image scaler is possible to include this sampling. The output of the upscaling-filtering cascade is an approximation of the continuous image signal, from which the correct

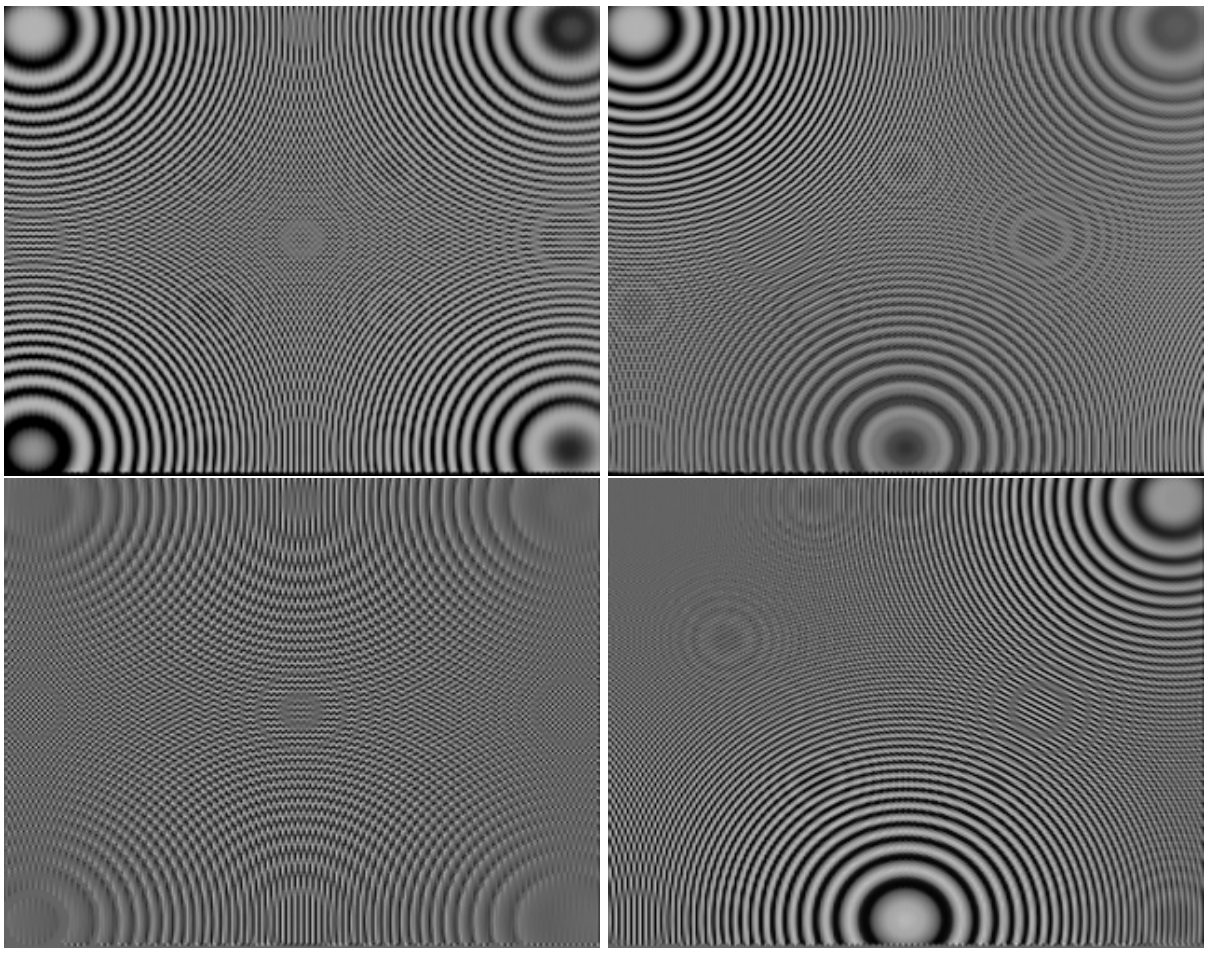


Figure 12: Zoneplate on a Delta-Nabla display. Left pixel sampling, right: subpixel sampling, top: luminance (Y), bottom: chrominance (U). With subpixel sampling, the spectrum repeats move to a diagonal high frequency, which increases the vertical resolution.

samples are taken to arrive at the desired resolution. Applying subpixel sampling simply means taking the right sample phases for R, G and B individually, and choosing a perceptually optimized filter. We shall call this 'subpixel scaling', and it is applicable for upscaling as well as downscaling.

Since the signal frequencies that profit most from the subpixel scaling are close to the Nyquist limit, the impact will be largest for downscaling applications. This is because the highest part of the baseband will be empty after upscaling.

5.1 Subpixel polyphase filtering

Polyphase filters have been proposed as an efficient implementation of a sampling rate converter [13, 6]. In a polyphase filter, samples not used after downsampling are not calculated, and multiplications with zeros from upsampling are omitted. When we further fix, either the factor K in case of upscaling, or the factor L in case of downscaling, a single design of the filter H, which depends on these factors, can be used. The polyphase structure can then be used with filter H for (almost) any up-, or downscaling factor, by varying the remaining factor (L or K). For upscaling, the low-pass filter should have a cut-off frequency near the Nyquist frequency of the input signal, but for downscaling it should be near the Nyquist frequency of the output signal to suppress aliasing. The use of only one filter for all (down)scaling factors enables us to achieve a simple color error versus sharpness trade-off control.

By separating horizontal and vertical scaling, 2D image scaling is reduced to two cascaded 1D sample rate converters.

Figure 14 shows how the polyphase filter is used for subpixel scaling: the relative phase, which corresponds

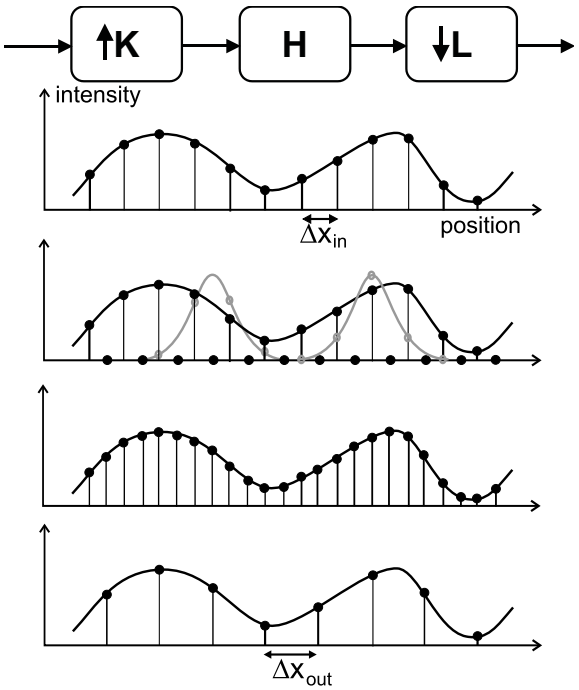


Figure 13: From top to bottom: The basic scheme of a sample rate converter, input signal, after upsampling, after filtering, after downsampling

to different coefficients in the filter, is set for the *RGB* signals corresponding to their relative position on the screen, and the scaling filter is chosen to optimize the trade-off between sharpness and color errors. The setting of this trade-off does not dependent on the scaling factor, because the filter is defined relative to the sampling grid of the display.

Figure 15 shows some different filters that can be used to set this trade-off. The exact choice of this trade-off will vary from display to display, for example with the (saturation of) the color points of the primary colors. We consider a more detailed discussion to be outside the scope of this paper.

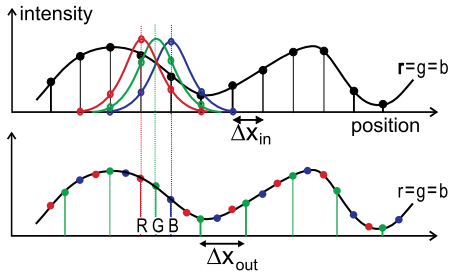


Figure 14: Subpixel scaling with a polyphase filter uses three different phases for RGB.

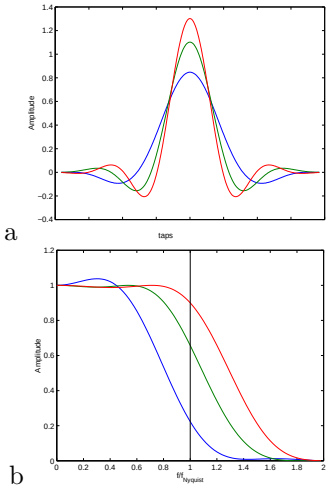


Figure 15: Filters with different color error - sharpness trade-offs. a) impulse response, b) frequency response.

The frequency responses of the filters also illustrate that it is difficult to separate the effects of subpixel sampling outside and inside the baseband (see also Sections 3.1 and 3.2). This is because practical filter responses will always have a finite transition band between pass- and stopband. This means that increasing the amplitude of frequencies just below the Nyquist limit will simultaneously increase the amount of aliasing. For this practical reason, the two effects described in sections 3.1 and 3.2 are strongly related.

The additional hardware cost of our proposed subpixel scaling is low and there is no longer a need for having an input signal at three times the resolution, since our subpixel scaling will work with any (down)scaling factor.

Subpixel scaling can be simply generalized for other subpixel arrangements. An issue is that the separation of a 2D scaler into two 1D scalers requires concessions on the horizontal/vertical or diagonal frequency response, as these reponses are no longer independent. Since for many subpixel structures the spectrum repeats occur in the diagonal part of the spectrum (see Figure 12), the diagonal frequency response becomes an important parameter too.

6 Results

Figures 16 and 17 show some results of subpixel downscaling on natural images and text. These figures are simulations of images on a display. Each subpixel is simulated using a number of pixels with only one primary color [14], but this results in relatively dark im-

ages. Therefore, the resulting images are low-pass filtered to include the effect of the human visual system in the simulation. This filtering also allows us to increase the brightness. We shall discuss the result for natural and text images in separate sub-sections.

6.1 Natural images

Figure 16 shows the results of downscaling a natural image with different scaling methods. Without subpixel scaling (Figure 16a,b), there is a trade-off between sharpness and aliasing, as can be seen by the increased jaggedness of the diagonal ropes in Figure 16b. When we turn to subpixel scaling, Figure 16a vs. Figure 16c shows an increased sharpness for comparable jaggedness, and Figure 16b vs. Figure 16c shows a decreased jaggedness for comparable sharpness. Finally, Figure 16d shows that the use of the two-dimensional Delta-Nabla subpixel arrangement also increases the sharpness in vertical direction, where the extra resolution for the vertical stripe arrangement only occurs in the horizontal direction.

6.2 Text and graphics

Subpixel scaling is also applicable to graphics, notably text. Previous methods [2, 8] were targeted at this category. These methods use oversampled rendering, and a fixed subpixel downscaling filter. This is usually referred to as 'anti-aliasing', but we prefer to look at it as downscaling an oversampled rendered image.

Text is particularly suited for improvement using subpixel techniques, not only because it contains very high frequencies, but also because traditional rendering methods suffer heavily from aliasing. A quality improvement by applying better anti-alias filtering, as in image and video processing, should be expected. Applying the proposed flexible subpixel scaling eliminates the need for a particular input resolution. Most of the quality improvement, which is also applicable to CRTs [18], is due to improved rendering at a higher resolution.

Figure 17 shows the results of subpixel scaling on a text image. The black and white text was not scaled, but directly rendered to the display resolution, a common practice in computer applications. Here, we clearly see the jaggedness that is caused by a total absense of anti-alias filtering, but which does result in very sharp edges. With anti-alias filtering, applied by downscaling oversampled text, the jaggedness decreases but remains clearly visible. This is a direct result of the imperfect reconstruction of the display (see Section 3.2). More

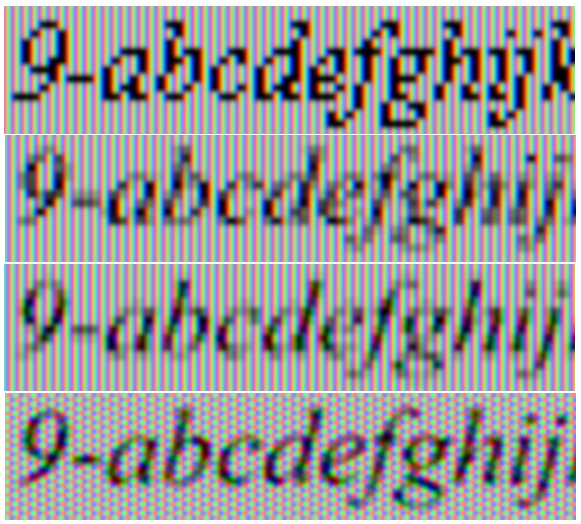


Figure 17: Subpixel scaling on text images. a) black and white text, b) with anti-alias filtering, c) with subpixel scaling, d) with subpixel scaling on a Delta-Nabla display

blurring is required to also suppress the distorted frequencies. This may be no problem on a CRT display that already suffers from blurring by the electron spot, but on a matrix display this is undesirable. With subpixel scaling, the jagginess is much reduced, and we can maintain a high level of sharpness without distortions. Obviously, it can never become as sharp as the black and white image, but this is the price to pay for an undistorted image. Finally, with the Delta-Nabla arrangement, there is an additional improvement in the vertical direction.

7 Conclusions

Subpixel image scaling is an attractive method to exploit the perceived resolution increase obtainable from the subpixel structure of matrix displays, particularly for downscaling applications. The method is applicable to any subpixel arrangement, poses no constraints on the input resolution, and is applicable to natural images as well as graphics and text rendered at a higher resolution.

A particular subpixel arrangment can only offer better perceived resolution than the vertical stripe arrangement, if the image resampling is adapted to this arrangement. This means that subpixel scaling is a crucial part of the signal processing chain if the display is supposed to offer resolution benefits with special subpixel arrangements.

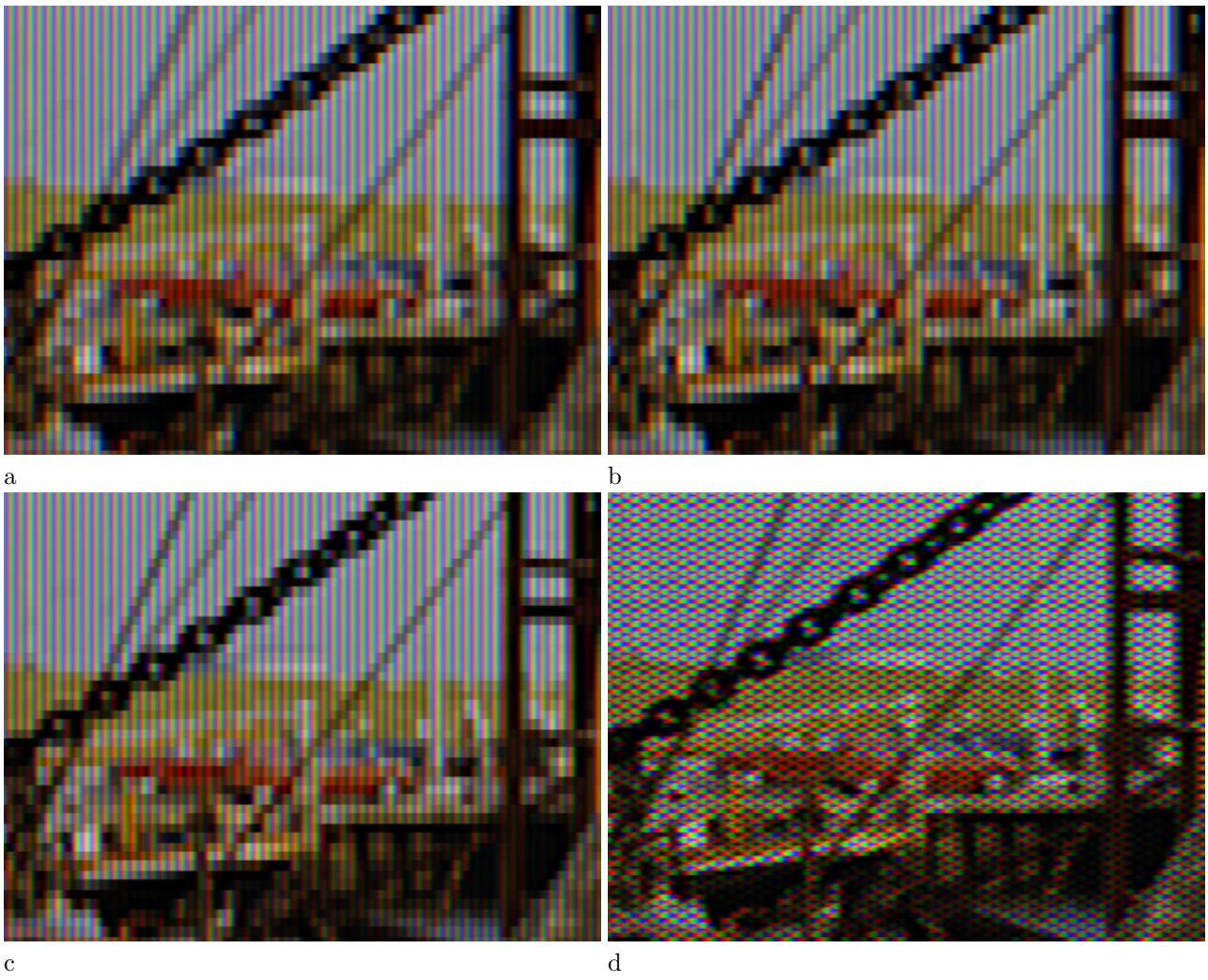


Figure 16: Subpixel scaling on (detail of) a natural image. a) unsharp & low jagginess, b) sharp & jagginess (e.g. in the diagonal ropes), c) subpixel scaling: sharp & low jagginess, d) subpixel scaling on a Delta-Nabla display

The increase in perceived resolution from subpixel scaling, does not imply that we can beat the sampling theorem. The Nyquist limit has not moved, i.e. the number of samples has not changed. The perceived resolution increase is caused by shifting luminance aliasing to the chrominance signal, resulting in decreased visibility of the distortions. This decreased visibility holds for frequencies below the Nyquist limit as well as for aliased components.

References

[1] Amano, Y., 'A Flat-Panel TV Display System in Monochrome and Color' *IEEE trans. on Electronic Devices*, Vol ED-22 no. 1, 1975, pp. 1–7.

[2] Betrisey, C., et. al., 'Displaced Filtering for Patterned Displays', *SID 2000 Digest*, pp. 296–299.

[3] Brown Elliott, C.H., 'Co-Optimization of Color AMLCD Subpixel Architecture and Rendering Algorithms', *SID 2002 Digest*, pp. 172–175.

[4] Chen, L.M., Hasegawa, S, 'Influence of pixel-structure noise on image resolution and color for matrix display devices', *J. SID 1/1*, 1993, pp. 103–110.

[5] Daly, S., 'Analysis of Subtriad Addressing Algorithms by Visual System Models', *SID 2001 Digest*, pp. 1200–1203.

[6] van den Enden, A., Verhoeckx, N.A.M., *'Discrete-time signal processing'*, Prentice Hall, ISBN 0132167638, 1989.

[7] Feigenblatt, R., 'Full color imaging on amplitude color mosaic displays', *Proc. SPIE*, Vol 1075, 1989, pp. 199–205.

[8] Gibson, S., 'Subpixel font rendering', *http://www.grc.com/cleartype.htm.*

[9] Hara, Z., Shiramatsu, N., 'Improvement in the picture quality of moving pictures for matrix displays', *J. SID 8/2*, 2000, pp. 129–137.

[10] Hsu, S.C., 'The Kell Factor, Past and Present', *SMPTE Journal*, February 1986. pp. 206–214.

[11] ITU, *recommendation ITU-R BT.601-5*, 1995

[12] Kranz, J.H., Silverstein, L.D., 'Color Matrix Display Image Quality: The Effects of Luminance and Spatial Sampling' *SID Digest*, 1990, pp. 29–32.

[13] Proakis, J.G., *'Advanced Digital Signal Processing'*, Macmillan Publishing Company, ISBN 0023968419, 1992

[14] Samadani, R., et. al., 'Periodic plane tilings: Application to pixel layout simulations for color flat-panel displays', *J. SID 2/2*, 1994, pp. 95–104.

[15] Silverstein, L.D., et. al., 'A psychophysical evaluation of pixel mosaics and gray-scale requirements for color matrix displays', *SID Digest*, 1989, pp. 128–131.

[16] Schr oder, H., Blume, H., *'One- and Multidimensional Signal Processing'*, John Wiley & Sons, ltd., ISBN 0471805416, 2000.

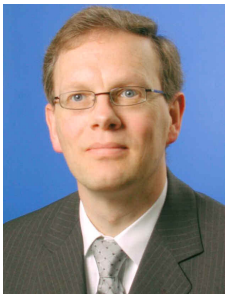
[17] Sluyterman, A.A.S., 'Resolution Aspects of the Shadow-Mask Pitch for TV Applications', *SID 1993 Digest*, pp. 340–343.

[18] Sluyterman, A.A.S., 'Improved Character Representation by Non-Square Pixel Grids', *SID 2000 Digest*, pp. 288–291.

[19] Wandell, B.A., *'Foundations of Vision'*, Sinauer Associates, Inc., ISBN 0878938532, 1995.

8 Biography

Michiel A. Klompenhouwer received the M.Sc. degree in Applied Physics from the University of Twente in 1998. He joined Philips research in 1998. Currently he is a Research Scientist in the group Video Processing and Visual Perception of Philips Research Laboratories. He has a particular interest in video processing for emerging television displays, and is working towards a Ph.D. degree in this field. He wrote a number of papers and his work resulted in several patent applications. In 2002, he received the Chester Sall Award from the IEEE Consumer Electronics Society. Michiel Klompenhouwer is a member of the SID.



Gerard de Haan received B.Sc., M.Sc., and Ph.D. degrees from Delft University of Technology in 1977, 1979 and 1992 respectively. He joined Philips Research in 1979. Currently he is a Research Fellow in the group Video Processing & Visual Perception of Philips Research Eindhoven and a Professor at the Eindhoven University of Technology. He has a particular interest in algorithms for motion estimation, scan rate conversion, and image enhancement. His work in these areas has resulted in several books, more than 80 scientific papers (www.ics.ele.tue.nl/˜dehaan/publications.html), around 50 patents and patent applications, and several commercially available ICs. He was the first place winner in the 1995 ICCE Outstanding Paper Awards program, the second place winner in 1997 and in 1998, and the 1998 recipient of the Gilles Holst Award. In 2002, he received the Chester Sall Award from the IEEE Consumer Electronics Society. The Philips 'Natural Motion Television' concept, based on his PhD-study received the European Innovation Award of the Year 95/96 from the European Imaging and Sound Association, its successor 'Digital Natural Motion' received a Wall Street Journal Europe Business Innovation Award 2001. Gerard de Haan is a Senior Member of the IEEE.