Multiband Pixel Colour Classification from HDMI Emissions

Dimitrije Erdeljan, Markus G. Kuhn
Department of Computer Science and Technology, University of Cambridge, UK
{de298, mgk25}@cl.cam.ac.uk

***Abstract*— We demonstrate methods to enhance the reconstruction of displayed information from the compromising emanations of HDMI or DVI video cables. Using a software-defined radio receiver, we acquire multiple recordings of such emissions for the same displayed image, at adjacent, overlapping reception bands. We first perform frequency alignment and coherent periodic averaging on each of these recordings individually. We then mutually align the resulting frames such that we can extract colour-identifying features for each displayed pixel across multiple reception bands. These features then go into a clustering algorithm to classify the signals emitted by different TMDS symbols. Finally, we build a graph data structure of the most common transitions between such symbols, and identify loops in this graph as candidates of pixel colours that cycle through multiple symbols due to the DC-balancing algorithm applied by the TMDS encoding. This can enhance the readability of eavesdropped text with some colour combinations, as we demonstrate for signals recorded at 12 metres distance.**

***Keywords*— TEMPEST, compromising emanations, eavesdropping, displays, software radio, information leakage, side-channel attacks.**

I. Introduction

The perhaps most well-known electromagnetic eavesdropping risk posed by unintentional radio-frequency emissions of computers, often referred to as TEMPEST, is the reconstruction of readable text shown on a raster display from the signals emitted by its video cable. From the earliest published demonstrations, by van Eck [1] in 1985, to later demonstrations on digital video interfaces [2], including with software-defined radio receivers [3], this usually involves simply passing the signal received through an amplitude-demodulating receiver, and then converting the resulting baseband signal into a raster image. The signal can then be improved by frame-periodic averaging. This works well for certain combinations of foreground and background colour, but can provide poor contrast for others [2]. Attempts to improve the eavesdropper's ability to obtain a good contrast for more colour combinations have included clustering of amplitudes [4], frequency demodulation [5] and quadrature-amplitude demodulation [6, Ch. 4], the latter more recently also combined with principal-component analysis [7]. Phase-coherent periodic averaging [7] can be used to improve the signal-to-noise ratio of signals acquired with software-defined radio receivers prior to demodulation into a display brightness or colour. This preserves at least two degrees of freedom of information per target pixel (the real and imaginary part of the complex-valued IQ samples received and averaged), and even more if we acquire several samples per target pixel.

However, there is another way to obtain more information about individual pixel colours, namely by changing the tuning frequency of the receiver. If we think of the waveform on the video cable as a sampled signal, with the pixel-clock frequency f_p as the sampling rate, then the spectral information would repeat itself across the radio spectrum after every f_p-wide frequency interval. This is because any signal sampled with frequency f in the time-domain repeats in the frequency domain with a period length f (and vice versa). In eavesdropping demonstrations, the receiver bandwidth B is usually chosen close to f_p, or a small multiple thereof. A significantly lower bandwidth would cause blurring of pixels (due to the $\approx B^{-1}$ impulse-response length of the band-selection filter), and a significantly higher bandwidth would risk including more interference signals.

If, however, we look at the video signal as a serial transmission of individual bits, at a bitrate f_b, which is ten times the pixel rate f_p on HDMI [8] and DVI [9] cables, then we can see that the spectrum actually repeats only after every $f_\mathrm{b} = 10 f_\mathrm{p}$-wide frequency interval. And if f_b is much larger than available receiver bandwidths, there will be multiple tuning frequencies that complement each other regarding the information available from a digital serial transmission. Our goal is, therefore, to combine into a single image information obtained from eavesdropping on multiple receiver bands. This requires techniques for carefully aligning images from multiple bands. After that, we can use a clustering algorithm to map the high-dimensional signal into an easy-to-distinguish set of false colours, in the hope of obtaining a good contrast for a wider set of combinations of foreground and background colours.

We also try to deal with another phenomenon that anyone having tried to eavesdrop on DVI or HDMI cables may be familiar with, namely the stateful behaviour of the *transition-minimized differential signalling* (TMDS) encoding [8], [9] used by both these interfaces. Many of the RGB byte values on the cable are not represented by a single 10-bit word, but chosen from two alternative 10-bit words, to maintain the DC balance of the transmitted signal [6, Appendix A]. While some 8-bit bytes are represented by a single *balanced* 10-bit word, which consists of five zero bits and five one bits, others have two possible 10-bit words that can represent them, including *unbalanced* ones with fewer or more zero bits than one bits.

The choice between the two is made by the encoder so as to maintain a balanced *running disparity* along each line of the video signal. This difference between the number of zero and one bits transmitted so far is kept close to zero by the encoder, and the count is reset to zero at the start of each line. As a result, even if a line consists entirely of a single repeated byte value, the encoder may go through a series of running-disparity states, resulting in a cycle of different 10-bit words. This is why uniformly coloured areas can show up to the eavesdropper as a series of vertical stripes or similar repeating patterns. Having already clustered the emissions from different 10-bit words, we, therefore, demonstrate here how an eavesdropper also can cluster cycles of 10-bit words, to reduce the visual impact of the TMDS DC-balancing algorithm on the appearance of the reconstructed false-colour raster image.

II. Signal Preparation

The initial signal acquisition and preprocessing steps, consisting of coherent demodulation and periodic averaging of IQ signals acquired with a software-defined radio receiver, are similar to those described in [7], therefore we will use similar notation. Readers may want to consult [7] for a more detailed explanation of some of those steps.

A. Software-defined Radio Reception

We use a software-defined radio (SDR) receiver to convert an analog antenna signal $s_{\rm a}(t)$ into a discrete sequence of samples $r[m]$ that can be recorded by a computer. We configure the SDR to acquire all radio signals within a frequency band $\left[f_{\rm c} - \frac{B}{2}, f_{\rm c} + \frac{B}{2}\right]$ around a centre frequency $f_{\rm c}$, with a sampling rate $f_{\rm r}$ that is somewhat greater than the desired bandwidth B.

The SDR then first downconverts the antenna signal, i.e. shifts its frequency by $-f_{\rm c}$, by multiplying with a phasor:

$$s_{\rm d}(t) = s_{\rm a}(t) \cdot {\rm e}^{-2\pi {\rm j} f_{\rm c} t}. \tag{1}$$

It then lowpass filters that signal with a cut-off frequency of $\pm\frac{B}{2}$, resulting in a band-limited complex-valued signal $s_{\rm f}(t)$, which it then samples into the discrete sequence

$$r[m] = s_{\rm f}(m/f_{\rm r}). \tag{2}$$

For multi-band experiments, we make multiple such recordings, at centre frequencies $f_{{\rm c},1} < f_{{\rm c},2} < \cdots < f_{{\rm c},n_{\rm b}}$. Where needed, we indicate the use of centre frequency $f_{{\rm c},i}$ by adding the corresponding band index i as a subscript, as in $r_i[m]$.

B. Resampling and Frequency Alignment

Having received the complex-valued IQ samples $r[m]$ from the SDR, we then need to perform two additional transforms before these can be turned into raster images or subjected to periodic averaging for noise reduction.

Firstly, we need to interpolate $r[m]$ back into $\tilde{s}_{\rm f}(t) \approx s_{\rm f}(t)$ such that we can resample it, to change the sampling frequency from $f_{\rm r}$ to the pixel-clock frequency $f_{\rm p}$ or some integer multiple $kf_{\rm p}$ of it. Reasonable choices for k are in the range $\lfloor B/f_{\rm p} \rfloor$ to $\lceil f_{\rm r}/f_{\rm p} \rceil$. This resampling ensures that each pixel, each line and each frame become represented by an integer number of IQ samples, which enables their conversion into a raster image without geometric distortions. This is also a prerequisite for later periodic averaging at the frame rate. (Linear or Lanczos interpolation work well for this.)

Secondly we need to shift the frequency spectrum of $\tilde{s}_{\rm f}(t)$ such that one of the harmonics of $f_{\rm p}$, which we will refer to as $hf_{\rm p}$, is shifted to 0 Hz. This compensates for the fact that when $r[m]$ was recorded, the SDR changed the phase of the samples representing each pixel, by multiplying with the complex phasor ${\rm e}^{-2\pi {\rm j} f_{\rm c} t}$. We now need to do this additional frequency shift in order to ensure that the samples associated with each pixel undergo the same phase change, such that we can perform coherent periodic averaging of complex values. Otherwise, they would cancel each other out due to the random phase imparted upon them by the local oscillator in the SDR and the pixel clock in the target device.

The result of both these transforms is a new discrete sequence of complex-valued IQ samples

$$s[n] \approx \tilde{s}_{\rm f}\left(\frac{n+\lambda}{f_{\rm s}}\right) \cdot {\rm e}^{2\pi {\rm j}(f_{\rm c} - hf_{\rm p})n/f_{\rm s}} \tag{3}$$

with new sampling rate $f_{\rm s} = kf_{\rm p}$. The offset λ can be adjusted to align $s[0]$ with the start of the next frame. The harmonic h of $f_{\rm p}$ can be chosen as $h = \lfloor f_{\rm c}/f_{\rm p} + 0.5 \rfloor$, the one nearest to $f_{\rm c}$.

However, for this to work, we need to know $f_{\rm p}$ with very high precision, with a relative error that is below around $0.1/(Nhkw_{\rm t}h_{\rm t})$, where N is the number of frames we want to average later and $w_{\rm t} \times h_{\rm t}$ is the total number of pixels of a frame in the video mode used, including blanking regions. This keeps the remaining phase rotation across N frames below a tenth of a full rotation.

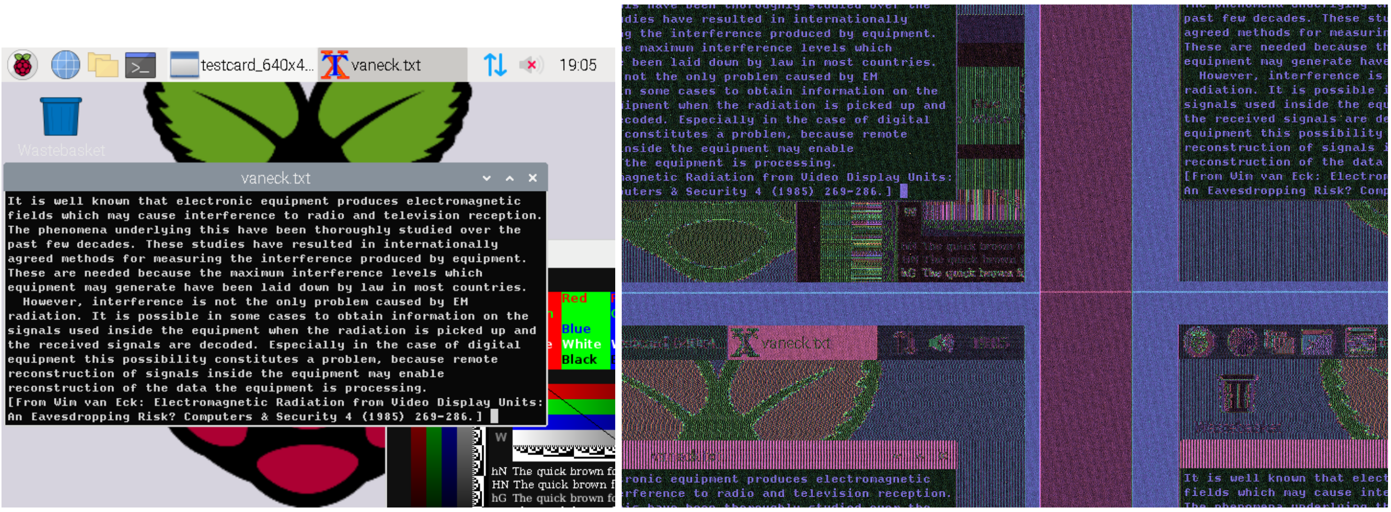
We exploit the periodicity of the video signal, which repeats with a frame rate of $f_{\rm v} = f_{\rm p}/(w_{\rm t}h_{\rm t})$ and look for a peak at lags of $f_{\rm v}^{-1}$ in the auto-correlation sequence of $s[n]$.

We iteratively refine our estimate for $f_{\rm p}$, starting from a nominal pixel frequency, as defined by a video-mode specification, such as [10], [11], or otherwise known for the target device (e.g., read from video-controller settings or measured with a spectrum analyzer). We start by generating an $s[n]$ sequence using equation (3) with this crude nominal $f_{\rm p}$ value. The length L of the generated sequence $s[n]$ must cover multiple frames, but should not be so long that $f_{\rm p}$ can be expected to drift by more than the aforementioned relative error. Otherwise subsequent frame averaging, in (7), of non-coherent complex values would reduce signal amplitude and increase blurring. In practice we use a couple of dozen frames.

We then estimate the auto-correlation sequence $R_{ss}[d] = \mathbb{E}\left[s[n]\, s[n+d]^*\right]$, for example using

$$R_{ss}[qa+d] \approx w^{-1} \sum_{n=0}^{w-1} s[n]\, s[n+qa+d]^* \tag{4}$$

for $d \in \{-d_{\max}, \ldots, d_{\max}\}$, where $s[0], \ldots, s[L-1]$ are the available samples, $q = kw_{\rm t}h_{\rm t}$ is the number of samples per



```
vaneck.txt
It is well known that electronic equipment produces electromagnetic
fields which may cause interference to radio and television reception.
The phenomena underlying this have been thoroughly studied over the
past few decades. These studies have resulted in internationally
agreed methods for measuring the interference produced by equipment.
These are needed because the maximum interference levels which
equipment may generate have been laid down by law in most countries.
  However, interference is not the only problem caused by EM
radiation. It is possible in some cases to obtain information on the
signals used inside the equipment when the radiation is picked up and
the received signals are decoded. Especially in the case of digital
equipment this possibility constitutes a problem, because remote
reconstruction of signals inside the equipment may enable
reconstruction of the data the equipment is processing.
[From Wim van Eck: Electromagnetic Radiation from Video Display Units:
An Eavesdropping Risk? Computers & Security 4 (1985) 269-286.]
```

Fig. 1. Displayed 640×480@60Hz test image (left), and eavesdropped 800×525-pixel frame $a[n]$ (right), recorded at $f_\mathrm{c} = 350$ MHz, after resampling, frequency alignment, periodic averaging of $N = 30$ frames, and HSV demodulation, but prior to frame alignment.

frame (that we want to converge to), a is the number of frame intervals over which we measure the autocorrelation, and $w \le L - qa - d_\mathrm{max}$ is the length of the correlation window used.

We then look for the peak

$$\hat{d} = \operatorname*{argmax}_{-d_\mathrm{max} \le d \le d_\mathrm{max}} |R_{ss}[qa + d]|^2 . \tag{5}$$

A peak location $\hat{d}$ means that a better frame-length estimate than q would have been $q + \hat{d}/a$. But rather than updating q by a factor $\frac{qa+\hat{d}}{qa}$, we instead divide f_p by that factor, to get the number of samples per frame closer to q at the next round of resampling.

In addition, if our estimate q of the number of samples per frame was, say, $\frac{1}{10}$ of a pixel wrong, then the phase angle $\angle R_{ss}[qa+\hat{d}]$ would show $\frac{ah}{10}$ of a rotation, or $\angle R_{ss}[aq+\hat{d}] = \frac{2\pi ah}{10}$, per time interval $(qa+\hat{d})/(kf_\mathrm{p})$. This is because in steps (1) and (3) we had frequency-shifted $s[n]$ such that the phasor applied to the antenna signal performs h rotations per pixel.

Combining both corrections, we update our estimate of the pixel-clock frequency based on the location and phase angle of the cross-correlation peak as

$$f_\mathrm{p} := f_\mathrm{p} \cdot \left(\frac{qa}{qa + \hat{d}} + \frac{k \angle R_{ss}[qa + \hat{d}]}{2\pi h(qa + \hat{d})} \right) . \tag{6}$$

After this update, we go back to resampling, and iterate steps (3), (4), (5), (6) a few times, until the process converges to $\hat{d} = 0$ and $|\angle R_{ss}[qa + \hat{d}]| < 10^{-5}$.

We now have in $s[n]$ a sequence of frames, each q samples long and with matching phase. Therefore, we can now perform coherent periodic averaging of up to $N = \lfloor L/q \rfloor$ frames, resulting in a complex-valued vector

$$a[n] = N^{-1} \sum_{l=0}^{N-1} s[n + ql], \ \text{for } n \in \{0, \ldots, q-1\} \tag{7}$$

We can rasterize $a[n]$ by reshaping it line-by-line into a matrix of size $kw_\mathrm{t} \times h_\mathrm{t}$, and converting each complex number into a colour pixel in HSV (hue, saturation, value) colour space, where the value (brightness) encodes the normalized absolute value $|a[n]|$ and the hue the angle $\angle a[n]$ of the complex number [7]. This results in false-colour raster images such as the one shown in Figure 1 (right).

While the above process for estimating f_p usually converges well to a frame length of q, it can occasionally result in frames being some $1/h$ of a pixel too long or short, which is recognizable in an HSV-demodulated averaged frame as one rotation of the hue along the height of the image. In such a case, we can manually fix the pixel-clock estimate as $f_\mathrm{p} := f_\mathrm{p} \cdot (1 \pm (hw_\mathrm{t} h_\mathrm{t})^{-1})$.

The accuracy of the f_p estimate can be somewhat improved by choosing the number a of frame lengths in the autocorrelation lag as more than 1, as this way the peak location $\hat{d}$ indicates the error in the frame length with a resolution of $1/a$ samples.

C. Inter-band Frame Alignment

To gain more information about the targeted signal, we acquire a sequence of SDR recordings $r_i[m]$, each taken at a different centre frequency $f_{\mathrm{c},i}$. We choose these SDR tuning frequencies such that the reception bands $[f_{\mathrm{c},i} - \frac{B}{2}, f_{\mathrm{c},i} + \frac{B}{2}]$ and $[f_{\mathrm{c},i+1} - \frac{B}{2}, f_{\mathrm{c},i+1} + \frac{B}{2}]$ overlap by several megahertz, i.e., such that $f_{\mathrm{c},i+1} - f_{\mathrm{c},i} < B$.

We first perform the signal preparation steps described in Section II-B independently for each of the recordings $r_i[m]$, resulting in averaged frame vectors $a_i[n]$. We treat $a_i[n]$ implicitly as periodic, that is $a_i[n] = a_i[n \bmod q]$.

Before attempting to classify pixel colours, we first need to align the averaged frames $a_i[n]$, such that a sample position n corresponds to the same displayed pixel in each of them. In some cases, the SDR may be able to measure the elapsed time between $r_i[0]$ and $r_{i+1}[0]$ with enough accuracy to allow some alignment, but the retuning step needed to change from $f_{\mathrm{c},i}$ to $f_{\mathrm{c},i+1}$ may disturb this timing information, and certainly

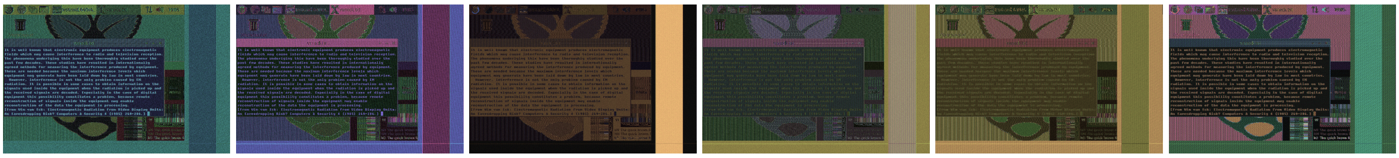


Fig. 2. The result of resampling, frequency aligning, periodically averaging (30 frames each), frame aligning and HSV demodulating six recordings acquired at a distance of 12 metres, at six centre frequencies, 25 MHz apart, from 325 MHz (top left) to 450 MHz (bottom right), with sampling frequency $f_\mathrm{r} = 64$ MHz.

the local-oscillator phase. Therefore, we assume here that such relative time information is not available from the SDR.

Instead, we exploit the overlap in our recordings $r_i[m]$ and $r_{i+1}[m]$ of adjacent frequency bands. After downconverting, resampling and averaging, frames $a_i[n]$ and $a_{i+1}[n]$ still contain some of the same information from the target's emissions, up to a frequency shift due to different centre frequencies, and so we can align them by correcting that frequency shift, filtering out only the frequencies contained in both reception bands, and then looking for a peak in the cross-correlation between these frames.

We recall that $a_i[n]$ had previously been frequency shifted such that the antenna frequency $h_i f_\mathrm{p}$ ended up at 0 Hz, while in $a_{i+1}[n]$ instead $h_{i+1} f_\mathrm{p}$ ended up at 0 Hz. Therefore, before we can compare both signals, we need to once more frequency-shift $a_{i+1}[n]$ by $(h_{i+1} - h_i) f_\mathrm{p}$, such that in both the same antenna frequency $h_i f_\mathrm{p}$ ends up at 0 Hz:

$$\begin{aligned} \tilde{a}_{i+1}[n] &= a_{i+1}[n] \cdot \mathrm{e}^{2\pi \mathrm{j} n (h_{i+1} f_\mathrm{p} - h_i f_\mathrm{p})/(k f_\mathrm{p})} \\ &= a_{i+1}[n] \cdot \mathrm{e}^{2\pi \mathrm{j} n (h_{i+1} - h_i)/k}. \end{aligned} \tag{8}$$

We still cannot yet correlate $a_i[n]$ and $\tilde{a}_{i+1}[n]$, since this frequency shift may cause the frequency spectrum of $a_{i+1}[n]$ to wrap around its $k f_\mathrm{p}$-wide period (since it is a sampled signal with sampling rate $k f_\mathrm{p}$). We avoid this by applying to both $a_i[n]$ and $\tilde{a}_{i+1}[n]$ a bandpass filter that lets through only frequencies in the shared interval $F_i = [(h_{i+1} - h_i) f_\mathrm{p} - \frac{B}{2}, \frac{B}{2}]$.

Since both averaged signals $a_i[n]$ and $\tilde{a}_{i+1}[n]$ are periodic with period length q, this filtering can conveniently be done in the frequency domain after applying to both the Discrete Fourier Transform (DFT)

$$A_i[v] = \sum_{n=0}^{q-1} a_i[n] \cdot \mathrm{e}^{-2\pi \mathrm{j} \frac{nv}{q}} \tag{9}$$

$$\tilde{A}_{i+1}[v] = \sum_{n=0}^{q-1} \tilde{a}_{i+1}[n] \cdot \mathrm{e}^{-2\pi \mathrm{j} \frac{nv}{q}} \tag{10}$$

This can be calculated efficiently via the Fast Fourier Transform, since q usually factors into many small prime numbers. The DFT of the (circular) cross-correlation of $a_i[n]$ and $\tilde{a}_{i+1}[n]$ is the product of $A_i[v]$ and $\tilde{A}_{i+1}[v]^*$. Therefore we apply the inverse DFT to obtain the cross-correlation sequence

$$R_{a_i \tilde{a}_{i+1}}[d] = \sum_{v=0}^{q-1} A_i[v] \cdot \tilde{A}_{i+1}[v]^* \cdot W_i[v] \cdot \mathrm{e}^{2\pi \mathrm{j} \frac{dv}{q}}. \tag{11}$$

The additional factor

$$W_i[v] = \begin{cases} 1, & \text{if } \mathrm{FFTfreq}(v, k f_\mathrm{p}) \in F_i \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

is the aforementioned frequency-domain bandpass filter with passband F_i. (In practice, the transition between 0 and 1 in $W_i[v]$ can be smoothed, e.g. using a raised-cosine roll-off.)

We then identify the amplitude peak

$$\Delta n = \operatorname*{argmax}_{d \in \{0, \ldots, q-1\}} \left| R_{a_i \tilde{a}_{i+1}}[d] \right|^2, \tag{13}$$

and align the two recordings by circularly shifting $a_{i+1}[n]$ by Δn.

For more than two recordings $a_1[n], a_2[n], \ldots, a_{n_\mathrm{b}}[n]$, we simply repeat this procedure for each adjacent pair. Fig. 2 shows $n_\mathrm{b} = 6$ frames aligned this way. For the first recording, we circularly shifted $a_1[n]$ to move the blanking intervals to near the edge of the frame. This can be done similarly, by cross-correlation with a manually-aligned reference frame in which the active pixels were replaced with zero. Alternatively the blanking intervals can also be recognized as rectangular regions of low variance.

III. TMDS Symbol Classification

After resampling to $k f_\mathrm{p}$, each displayed pixel corresponds to k consecutive samples in each averaged frame $a_i[n]$. We assume here that pixel boundaries line up with sample boundaries, such that e.g. samples $a_i[0], a_i[1], \ldots, a_i[k-1]$ comprise one pixel. In practice, dividing $a_i[n]$ into k-sample-long segments will likely result in each segment partially covering two pixels, resulting in misclassified pixels on the boundary between two TMDS symbols. The eavesdropper can work around this limitation by trying several images for different cyclic shifts of $a[n]$, for example using a fractional delay filter, or by varying λ in (3) over a range $0 \leq \lambda < k$.

Since we have frequency-shifted in each recording a multiple $h f_\mathrm{p}$ of the pixel clock to 0 Hz, each ten-bit TMDS symbol will show up in $a_i[n]$ as the same sequence of k complex samples (plus noise) regardless of its position in the image. This is because we multiplied the received signal $s_\mathrm{a}(t)$ by a phasor $\mathrm{e}^{-2\pi \mathrm{j} h f_\mathrm{p} t}$, which during a k-samples long pixel period rotates by $-2\pi h f_\mathrm{p} \cdot f_\mathrm{p}^{-1} = -2\pi h$ radians. Since h is an integer, this means that there is no phase change between pixels due to downconversion.

Additionally, if we have several aligned frames $a_1[n], a_2[n], \ldots, a_{n_\mathrm{b}}[n]$ created by processing recordings made at different SDR centre frequencies, we expect that a TMDS symbol will result in matching k sample long

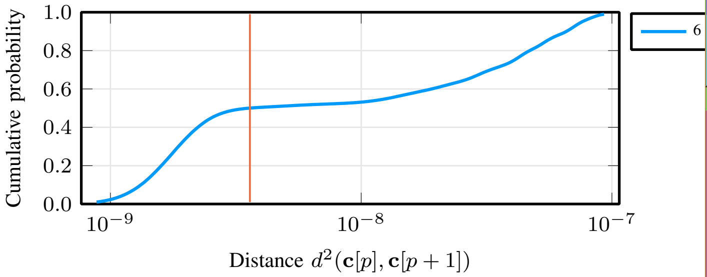


Fig. 3. Distribution of distance $d^2(\mathbf{c}[p], \mathbf{c}[p+1])$ between pairs of horizontally adjacent pixels, for pixel vectors obtained in our demonstration.

sequences within each of these frames. The p-th pixel then corresponds to a feature vector $\mathbf{c}[p] \in \mathbb{C}^{kn_\mathrm{b}}$:

$$\begin{aligned}\mathbf{c}[p] = (a_1[kp+0], a_1[kp+1], \ldots,\ \ & a_1[kp+(k-1)], \\ a_2[kp+0], a_2[kp+1], \ldots, a_{n_\mathrm{b}}&[kp+(k-1)]).\end{aligned} \tag{14}$$

We use the squared L^2 norm

$$d^2(\mathbf{c}[p_1], \mathbf{c}[p_2]) = (\mathbf{c}[p_1] - \mathbf{c}[p_2])^*(\mathbf{c}[p_1] - \mathbf{c}[p_2]) \tag{15}$$

as the distance between the feature vectors for pixels p_1 and p_2, where * denotes the conjugate transpose.

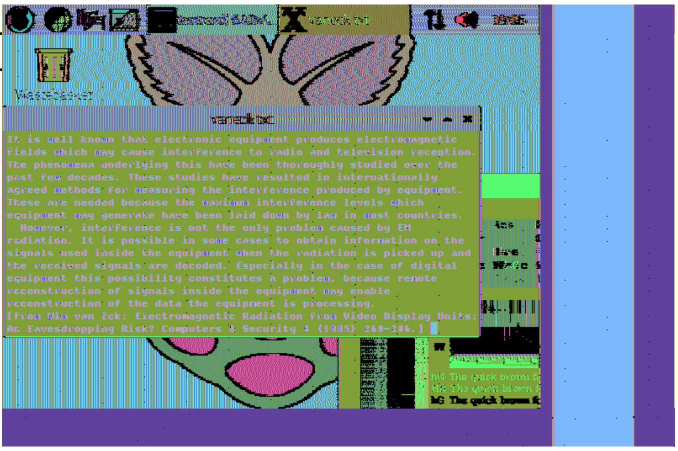
Figure 3 shows, for our example recordings, the cumulative distribution of distances $d^2(\mathbf{c}[p], \mathbf{c}[p+1])$ between pairs of horizontally adjacent pixels. If we use all $n_\mathrm{b} = 6$ recordings, resulting in $kn_\mathrm{b} = 3 \times 6 = 18$-dimensional feature vectors, this results in the thick blue line, which shows a clear bimodal distribution. From this plot, we choose a threshold $\varepsilon = 0.6 \times 10^{-4}$, and assume that TMDS symbols at positions p_1 and p_2 are the same if

$$d^2(\mathbf{c}[p_1], \mathbf{c}[p_2]) < \varepsilon^2. \tag{16}$$

In our example recordings, combining information from multiple bands significantly improved the separation between matching and differing TMDS symbol pairs. For comparison, if we plot the same distribution for each individual frequency band (the thin black lines), most lack this clear separation. On the other hand, using only the information from three non-overlapping frequency bands (green lines), works almost as well as using all six, even if none of them individually showed a bimodal distribution.

Next, our TMDS symbol classification algorithm outputs a label $C[p]$ for each pixel, where $C[p_1] = C[p_2]$ for all pixel positions p_1 and p_2 with $d^2(\mathbf{c}[p_1], \mathbf{c}[p_2]) < \varepsilon^2$. In other words, it finds the set of equivalence classes of pixel positions for an equivalence relation that is the transitive closure of (16). This can simply be computed by initialising all $C[p]$ to different values, and then merging the labels of each pair p_1, p_2 for which (16) holds. (For performance, we approximate this by first only considering p_2 that are in a 40-pixel square centered at p_1, and then merge the resulting clusters by comparing their average pixel vectors.)

Finally, we visualize these labels as a false-colour image by assigning a different random colour to all pixels with the same TMDS symbol $C[p]$, except those where $C[p]$ appears in fewer than 20 pixels, which we display as black. An example result can be seen in Figure 4 and magnified in Figure 6 (top).



It is well known that electronic equipment produces electromagnetic fields which may cause interference to radio and television reception. The phenomena underlying this have been thoroughly studied over the past few decades. These studies have resulted in internationally agreed methods for measuring the interference produced by equipment. These are needed because the maximum interference levels which equipment may generate have been laid down by law in most countries.

However, interference is not the only problem caused by EM radiation. It is possible in some cases to obtain information on the signals used inside the equipment when the radiation is picked up and the received signals are decoded. Especially in the case of digital equipment this possibility constitutes a problem, because remote reconstruction of signals inside the equipment may enable reconstruction of the data the equipment is processing.

[From Wim van Eck: Electromagnetic Radiation from Video Display Units: An Eavesdropping Risk? Computers & Security 4 (1985) 269-286.]

Fig. 4. The result of clustering the feature-vector space, where each colour indicates a label assigned to a likely TMDS symbol. Areas emitting balanced TMDS symbols appear uniform, but other areas show patterns of alternating TMDS symbols used for DC balancing.

IV. TMDS Cycle Detection

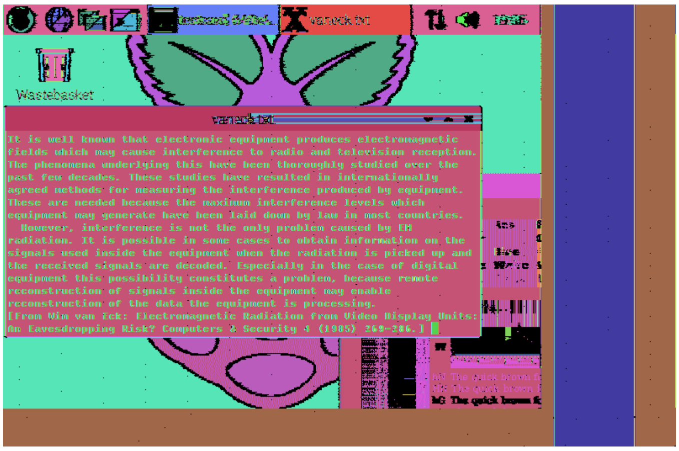
Since the TMDS encoding used in HDMI may have two possible ten-bit encodings for each red, green or blue byte, pixels with the same colour can appear as different TMDS symbols in the output of our classifier. Although there are at most two possible encodings for a byte, an RGB colour can appear as up to 2^3 different feature vectors $\mathbf{c}[p]$, since we receive a combination of the emissions of the red, green and blue lane, each of which uses a separate encoder. A single-colour region will appear as a periodic sequence of labels $C[p], C[p+1], \ldots$ as the TMDS encoders cycle through states (running disparities).

In order to identify TMDS symbols $C[p]$ that correspond to the same RGB colour, we search for sequences of symbols that are likely to appear consecutively. More formally, let $K(C_1, C_2)$ be the count of how often TMDS symbol label C_2 follows (immediately to the right of) symbol label C_1 in the frame. We consider C_2 to be a likely successor of C_1 in these encoder cycles if C_2 appears at least a fraction $\eta = 0.3$ as often after C_1 as the most frequent successor of C_1. We define a set E of such likely successor symbols as

$$E = \{(C_1, C_2) : K(C_1, C_2) \geq \eta \max_{C'} K(C_1, C')\}. \tag{17}$$

We expect that commonly repeated sequences of TMDS symbols show up as cycles in the graph $G = (C, E)$, the nodes of which are TMDS symbol labels, with directed edges between likely successor symbols. We, therefore, merge the labels C_1 and C_2 as different encodings of the same RGB colour if there is in G both a path from C_1 to C_2 and a path from C_2 to C_1, which can be computed efficiently using Tarjan's strongly connected components algorithm.

Merging strongly-connected labels this way for our demonstration resulted in Figure 5 (magnified in Figure 6 bottom).



```
It is well known that electronic equipment produces electromagnetic
fields which may cause interference to radio and television reception.
The phenomena underlying this have been thoroughly studied over the
past few decades. These studies have resulted in internationally
agreed methods for measuring the interference produced by equipment.
These are needed because the maximum interference levels which
equipment may generate have been laid down by law in most countries.
  However, interference is not the only problem caused by EM
radiation. It is possible in some cases to obtain information on the
signals used inside the equipment when the radiation is picked up and
the received signals are decoded. Especially in the case of digital
equipment this possibility constitutes a problem, because remote
reconstruction of signals inside the equipment may enable
reconstruction of the data the equipment is processing.
[From Wim van Eck: Electromagnetic Radiation from Video Display Units:
An Eavesdropping Risk? Computers & Security 4 (1985) 269-286.]
```

Fig. 5. The result of cycle detection. Note how the words "vaneck.txt" and "Wastebasket" become readable, which in Fig. 4 remained unrecognizable within the surrounding TMDS-balancing cycles.

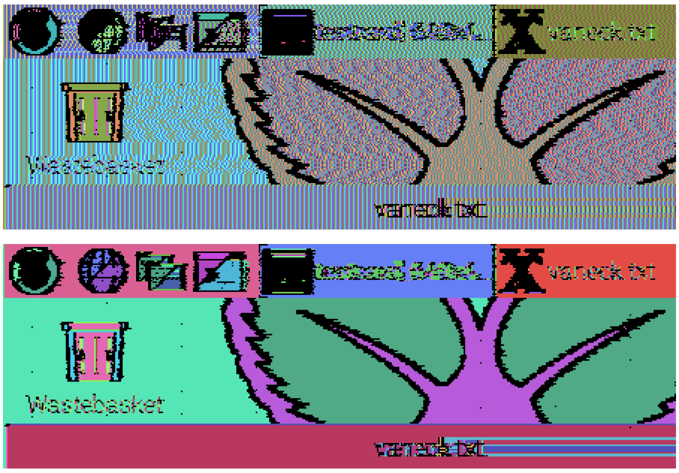


Fig. 6. Magnified versions of the output frames after TMDS symbol classification (top) and cycle merging (bottom).

V. Conclusion

While the results remain far from perfect, we have demonstrated that by combining spectral information from multiple bands, we can cluster the IQ values associated with individual TMDS symbols. This move from a continuous high-dimensional signal to a discrete set of symbols then allows us to build a graph that very roughly models the state transitions that occur in the TMDS encoder. Merging cyclic symbol sequences then results in a new classification that better represents uniform colour areas, such as text background, which can help to make foreground text visible that previously had remained hidden in the patterns caused by the TMDS DC balancing algorithm.

Appendix

Demonstration Setup

The target system in our demonstration consisted of a Raspberry Pi Model B+ computer connected via a 2.0 m long HDMI-to-DVI cable to a Dell 1704FPT display, operating in a 640×480@60 Hz mode, with a pixel-clock frequency of 25.2 MHz. The receiver consisted of a log-periodic antenna (Schwarzbeck VUSLP 9111B, 200–3000 MHz) connected to a Rohde & Schwarz FSV7 signal analyzer operating in IQ mode, with a flat bandwidth of $B = 40$ MHz and a sampling frequency of $f_\mathrm{r} = 64$ MHz. The recordings took place at 12 m distance in the corridor of a normal, unshielded office building.

We used $n_\mathrm{b} = 6$ recordings centered at $f_\mathrm{c} = 325, \ldots, 450$ MHz, in steps of 25 MHz, resampled to $f_\mathrm{s} = kf_\mathrm{p}$ with $k = 3$, before averaging $N = 30$ frames of $q = kw_\mathrm{t}h_\mathrm{t} = 3\times800\times525$ samples each (including blanking regions).

The recordings and Julia code for the demonstration shown in the figures are available at

https://www.cl.cam.ac.uk/research/security/datasets/hdmi/

References

[1] W. van Eck, "Electromagnetic radiation from video display units: An eavesdropping risk?" *Computers & Security*, vol. 4, no. 4, pp. 269–286, 1985, ISSN: 0167-4048. DOI: 10.1016/0167-4048(85)90046-X.

[2] M. G. Kuhn, "Electromagnetic eavesdropping risks of flat-panel displays," in *Privacy Enhancing Technologies*, D. Martin and A. Serjantov, Eds., ser. Lecture Notes in Computer Science, Berlin, Heidelberg: Springer, 2005, pp. 88–107. DOI: 10.1007/11423409_7.

[3] M. Marinov, "Remote video eavesdropping using a software-defined radio platform," Master's thesis, University of Cambridge, Computer Laboratory, Jun. 2014. [Online]. Available: https://github.com/martinmarinov/TempestSDR/.

[4] P. De Meulemeester, B. Scheers, and G. A. Vandenbosch, "Reconstructing video images in color exploiting compromising video emanations," in *2020 International Symposium on Electromagnetic Compatibility – EMC EUROPE*, ISSN: 2325-0364, Sep. 2020, pp. 1–6. DOI: 10.1109/EMCEUROPE48519.2020.9245775.

[5] ——, "Differential signaling compromises video information security through AM and FM leakage emissions," *IEEE Transactions on Electromagnetic Compatibility*, vol. 62, no. 6, pp. 2376–2385, Dec. 2020, ISSN: 1558-187X. DOI: 10.1109/TEMC.2020.3000830.

[6] C. D. O'Connell, "Exploiting quasiperiodic electromagnetic radiation using software-defined radio," PhD thesis, University of Cambridge, Computer Laboratory, 2019. [Online]. Available: https://www.repository.cam.ac.uk/handle/1810/290902.

[7] D. Erdeljan and M. G. Kuhn, "Benfits of coherent demodulation for eavesdropping on HDMI emissions," in *International Symposium on Electromagnetic Compatibility – EMC Europe*, IEEE, 2024, pp. 263–268. DOI: 10.1109/EMCEurope59828.2024.10722379.

[8] *High-Definition Multimedia Interface – Specification 1.3a*, 2006.

[9] *Digital Visual Interface – DVI, Revision 1.0*, Digital Display Working Group, 1999. [Online]. Available: https://glenwing.github.io/docs/.

[10] *VESA and Industry Standards and Guidelines for Computer Display Monitor Timing (DMT), Version 1.0, Revision 13*, Video Electronics Standards Association (VESA), 2013.

[11] *Coordinated Video Timings (CVT) Standard, Version 2.1*, Video Electronics Standards Association (VESA), 2023.