



Research & Development
White Paper

WHP 230

August 2012

**Measurement of Human Sensitivity across
the Vertical-Temporal Video Spectrum
for Interlacing Filter Specification**

K.C. Noland

BRITISH BROADCASTING CORPORATION

Measurement of Human Sensitivity across the Vertical-Temporal Video Spectrum for Interlacing Filter Specification

Katy C. Noland

Abstract

Good quality conversion from progressive to interlaced video is highly relevant to today's broadcast systems, in which interlaced content is still common. The interlacing process is a form of downsampling, and hence requires an anti-alias filter. For best results the anti-alias filter should be matched to the reconstruction filter, which is comprised of the display and the human visual system. Additionally, it must meet the technical requirements of the downsampling process. In this paper we present a novel method of measuring the combined response to interlacing artefacts that is simple and powerful. We use the results to derive an optimal anti-alias filter template, using a new region-growing technique that is specifically designed to match the measured response whilst keeping to the technical constraints of an interlaced sampling structure. Our results provide support for an existing, heuristically-defined filter, and show that the same filter could be used for a range of viewing distances.

This article was originally published in the Proceedings of the IEEE International Conference on Multimedia and Expo (ICME), Melbourne, Australia, July 2012. © IEEE.

Additional key words: interlacing, standards conversion, aliasing, filter design, human visual system

White Papers are distributed freely on request.
Authorisation of the Controller of Research & Development
is required for publication.

©BBC 2012. Except as provided below, no part of this document may be reproduced in any material form (including photocopying or storing it in any medium by electronic means) without the prior written permission of BBC Research & Development except in accordance with the provisions of the (UK) Copyright, Designs and Patents Act 1988.

The BBC grants permission to individuals and organisations to make copies of the entire document (including this copyright notice) for their own internal use. No copies of this document may be published, distributed or made available to third parties whether by paper, electronic or other means without the BBC's prior written permission. Where necessary, third parties should be directed to the relevant page on BBC's website at <http://www.bbc.co.uk/rd/pubs/whp> for a copy of this document.

Measurement of Human Sensitivity across the Vertical-Temporal Video Spectrum for Interlacing Filter Specification

Katy C. Noland

1 Introduction

Interlacing is commonplace in broadcast systems across the world [1]. It serves as a simple means of 2:1 data compression whilst maintaining a high refresh rate and vertical detail. This is achieved by omitting all odd lines of one progressive frame, and all even lines of the next, to form two fields that together maintain the original vertical resolution. The data loss is manifested as *interlace twitter*, an unnatural motion of vertical detail that is related to, but different from, an object's movement, caused by the time offset between fields.

Interlace twitter can be analysed as aliasing distortion resulting from the downsampling operation of omitting picture lines [2], and as such it can be minimised by preceding the downsampling with a suitable anti-alias filter to remove any content at frequencies that cannot be supported by the new sampling rate. Interlacing cannot be separated into individual vertical and temporal downsampling steps: both dimensions must be considered together. This means that the ideal antialias filter must be two-dimensional, and hence there is a range of possible two-dimensional shapes for the filter passband boundary, making the ideal filter non-trivial to define.

This problem is highly-relevant to today's broadcast systems, which still use interlacing extensively, especially for genres such as sport that have a lot of fast motion. Production environments are moving towards high definition (HD), high frame rate signal chains, but the data rates at transmission are still restricted. There is therefore a need to achieve the best possible interlacing performance at a converter just before coding and transmission, in order to reap the greatest benefit from improved production tools.

In this study we have for the first time directly measured the complete reconstruction filter response, which comprises the human sensitivity to interlacing artefacts across the vertical-temporal frequency spectrum and the response of the display. We used a liquid crystal display (LCD), the typical technology in use in today's homes. We have used this data to derive an ideal anti-alias filter template for interlacing, to suppress the signal components that produce the most visible artefacts. We target the specific case of converting progressive (non-interlaced) video of frame size 1920×1080 pixels and running at 50 frames per second, 1080p50, to interlaced video of the same frame size running at 50 fields per second (equivalent to 25 frames per second), 1080i25.

In section 2 we present the principles and constraints of interlacing filter design, and describe related work. In section 3 we introduce a new objective test paradigm that we have developed specifically for the task of measuring responses to interlacing artefacts. We also present a novel technique for deriving the ideal filter response from our measured data, that takes into account the constraints on the response shape imposed by interlacing theory. Our results are presented and discussed in section 4, and we summarise our findings in section 5.

2 Principles and Related Work

Although interlacing has been used since the earliest days of television, analysis is limited in the first publications on the topic, and filtering is not offered as a solution to observed errors [3] [4, p. 566]. Interlacing largely worked because there was low-pass filtering intrinsic to the system, in camera

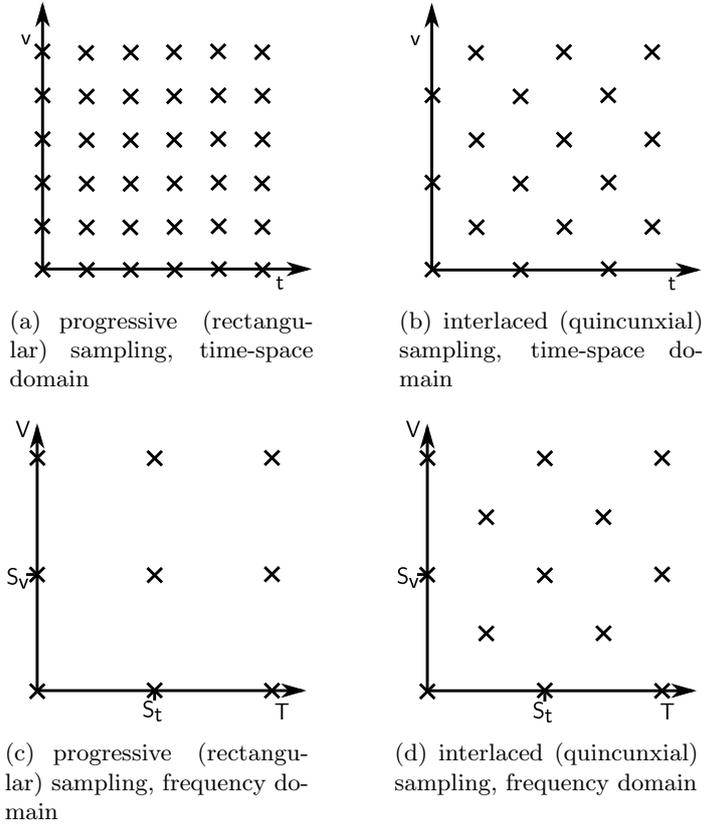


Figure 1: Vertical-temporal sampling structure in the time-space domain for (a) progressive and (b) interlaced pictures, and their respective frequency-domain equivalents (c) and (d). Vertical spatial position is denoted by v and time by t . V and T represent vertical and temporal frequencies, with S_v and S_t the progressive vertical and temporal sampling rates respectively. The horizontal spatial dimension is separable and can be imagined extending into the page.

apertures and integration times, in the scanning spot profile and spatial and temporal response of the phosphor in cathode ray tube displays, and in the human visual response.

It was not until the 1980s that a frequency-domain analysis of interlacing in the context of classical sampling theory was performed [1] and the need for optimised filtering was put forward. Interlacing video can be viewed as a two-dimensional downsampling problem, since the horizontal spatial dimension is not affected by the process. Progressive video is sampled on a rectangular lattice in vertical space and time, as shown in figure 1(a), whereas interlaced video is sampled on a quincunxial lattice (figure 1(b)). When observed in the frequency domain, any sampled signal consists of the original baseband spectrum plus aliases of the baseband centred on points of the reciprocal lattice, which is obtained by performing the Fourier transform on the sampling lattice. The reciprocal lattice is rectangular for progressive video (figure 1(c)) and quincunxial for interlaced video (figure 1(d)). Dubois [2] gives a rigorous explanation of sampling theory on a range of lattice structures.

In order for the aliases not to overlap, the signal must be filtered prior to sampling such that the baseband spectrum fills only a *unit cell* of the reciprocal lattice. This is one fundamental unit of area for the lattice, each one associated with a lattice point. Unit cells are of fixed size and shape such that all unit cells span the lattice area but do not overlap, i.e. they tessellate. This will be an important constraint for our interlacing filter. Using real signals adds the constraint that unit cells must be symmetrical about both sampling axes (vertical and temporal), hence each lattice point must be in the centre of its unit cell. Figure 2 shows some possible unit cells on rectangular (2(a)) and quincunxial (2(b)–2(d)) lattices.

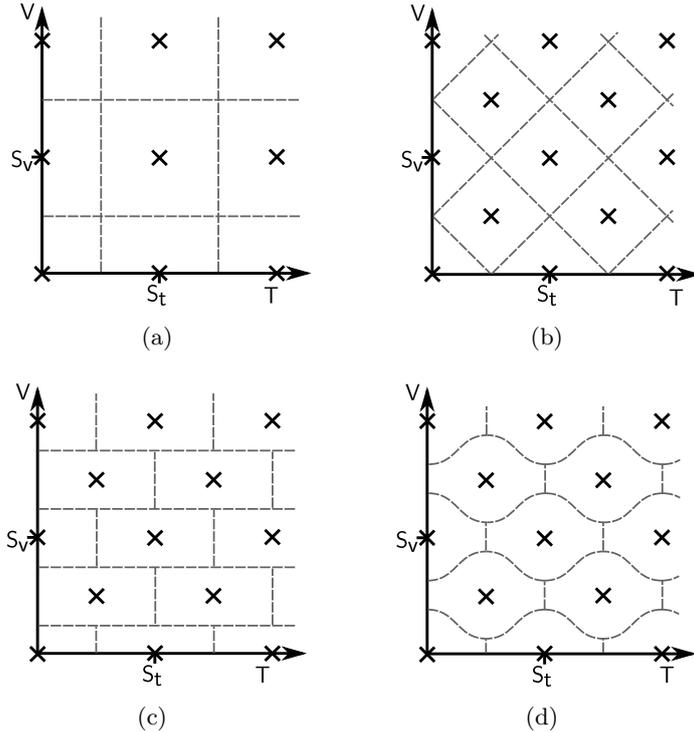


Figure 2: Possible unit cells for progressive video (a) and interlaced video (b), (c) and (d). S_v and S_t are the progressive vertical and temporal sampling rates respectively.

The *Voronoi cell*, or *Brillouin zone*, the region in which every point is closer to its associated lattice point than any other, is often assumed to be the optimum unit cell [1, 5, 6]. Figures 2(a) and 2(b) show the Voronoi cells for progressive and interlaced video respectively, assuming equivalence between the vertical and temporal sampling frequencies. This assumption, however, is likely to be invalid, since space and time are measured in fundamentally different units. Any scaling of the axes with respect to each other to make the subjective distances in both dimensions equivalent would result in a different Voronoi cell [7, p. 48]. Using the Voronoi cell of figure 2(b) as the target for an interlacing filter results in blurring of fast moving objects with any vertical detail, since most energy at high temporal frequencies is rejected.

An alternative unit cell for interlaced video that is not a Voronoi cell, achieved by pure vertical filtering, is shown in figure 2(c). This is the approach for interlaced output used in many professional cameras, which simply apply a two-line moving average vertical filter [8, p. 57]. This filter is also non-ideal, since it sacrifices vertical detail above a quarter of the vertical sampling rate even for stationary objects, and it has a very slow roll-off, which means that aliases are poorly suppressed.

It is clearly not sufficient to approach interlacing from a purely technical point of view: it is necessary to frame the problem in the context of its intended application, that of human viewing. The approach illustrated in figure 2(d) approximates the magnitude response of a filter proposed for de-interlacing [9], which is also of interest as a potential pre-interlacing filter since in an ideal system the pre- and post-interlacing filters should be matched [10]. It was developed heuristically [10], but is the first to take account of the human visual system as part of the filter design process. It allows more spatial detail to pass at high temporal frequencies than 2(b), at the expense of some spatial detail in slow-moving objects.

In the same study [10] Weston and Ackroyd also developed a filter template that resembles figure 2(b), derived from measured human contrast sensitivity data. Interestingly, when realised as a 3-field, 10-line implementation, the straight edges of the filter template become curved and the response is similar to figure 2(d). This is the closest approach to ours, although the measured

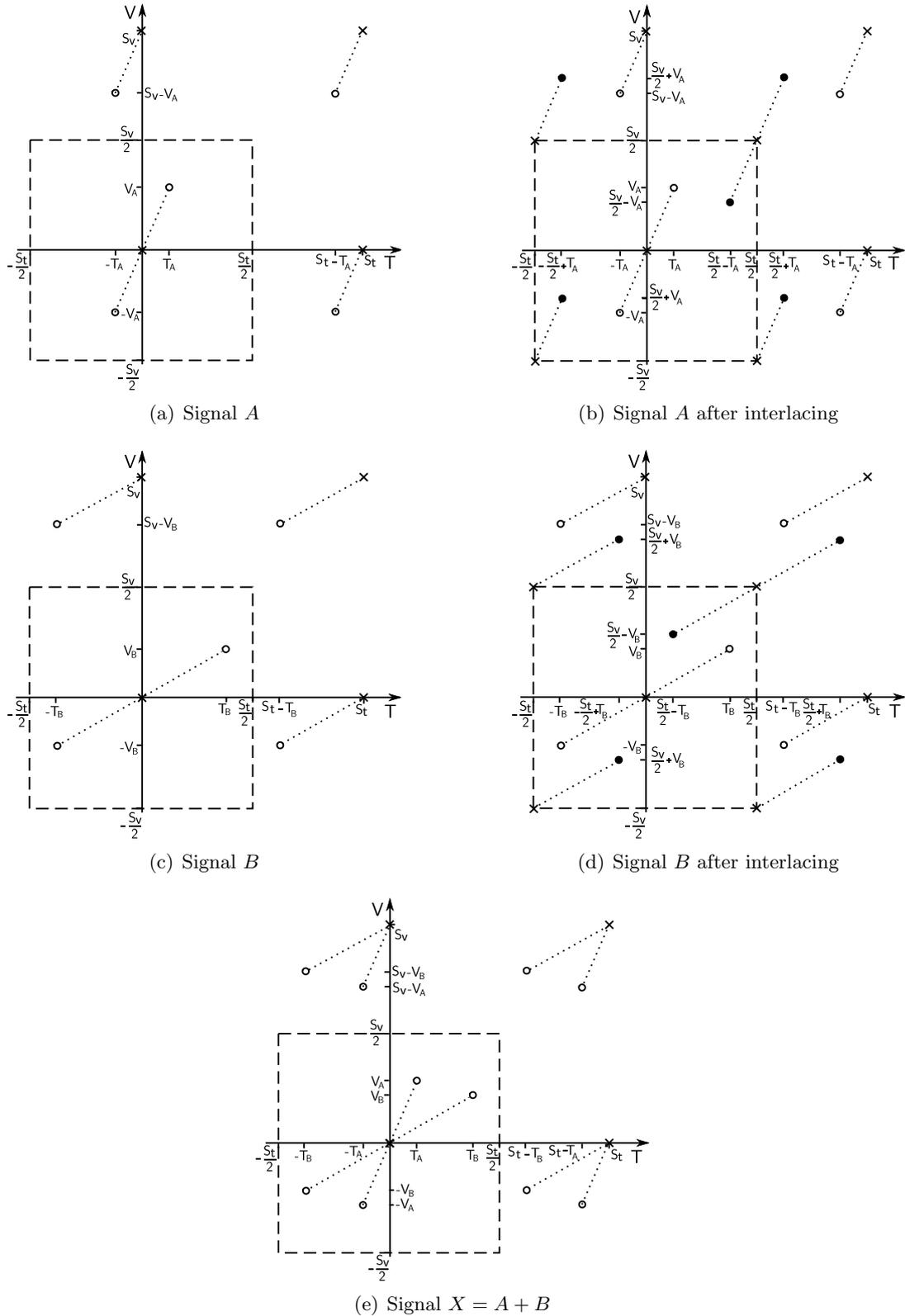


Figure 3: Example of the test signals used for the experiments, shown in the frequency domain. S_v and S_t are the vertical and temporal sampling rates respectively, and V_A and T_A are the vertical and temporal frequencies of signal A. Crosses indicate points of the sampling lattice, hollow circles are baseband signals and filled circles are aliased signals. The dashed square borders the progressive baseband spectrum.

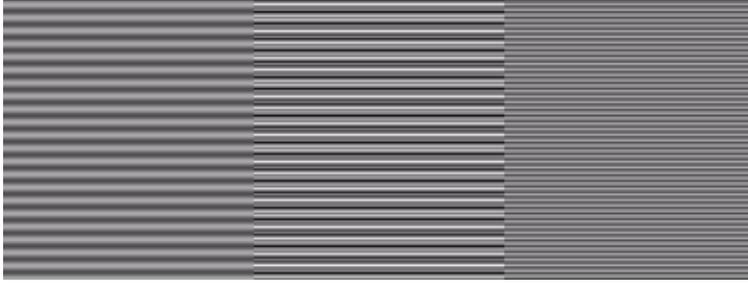


Figure 4: Example of the test signals used for the experiments, from left to right, A , X and B . In this case $V_A = 0.1471S_V$ and $V_B = 0.3529S_V$.

data represents human sensitivity to individual sinusoids whereas we concentrate on measuring perception of interlaced aliases.

Research on interlacing methods has been largely neglected in recent years, although the potential of linear vertical-temporal filtering has not yet been fully explored. There is, however, a large body of work that focusses on the process of de-interlacing. Suggested techniques include linear vertical-temporal filtering, motion adaptation, motion compensation, and hybrids of these. An excellent overview is given by de Haan and Bellers [6], with additional recent advances reviewed by Keller et al. [11]. Motion-adaptive and motion-compensated approaches to de-interlacing promise good quality conversion, at the expense of high complexity.

In this paper we return to the interlacing process and linear filtering, and conduct the first direct measurement of the perceptibility of interlacing artefacts. From our measurements we derive a perceptually-optimised linear vertical-temporal filter template. A filter based on this template would produce the best possible quality interlaced video for viewing on interlaced displays, and, since maximum conversion efficiency can only be achieved if all filters in the chain are matched [10, 6], it would also serve as the best quality linear de-interlacer for the same material.

3 Experiment Design

Previous experiments that have measured the human spatio-temporal contrast sensitivity function [12, 13] required subjects to judge the threshold of visibility for single sinusoids or square-waves. We take an alternative approach of pairwise comparisons of sinusoids, that is much simpler for the subjects, and ideally suited to our goal of measuring the visibility of interlacing artefacts. We are interested in the response when the eye is free to move, as is the case when watching television, so we also do not require any intrusive eye stabilisation methods [13]. Our experiment addressed the conversion of 1080p50 material to 1080i25, but the method is not format-specific. We refer to the original progressive vertical sampling frequency of 1080 lines per picture height as S_v , and the original progressive temporal sampling rate of 50 Hz as S_t .

Figure 3(a) shows the supported spectrum for 1080p50, with a vertical-temporal sinusoid A at frequency (T_A, V_A) . The process of interlacing introduces an extra alias into the square progressive baseband spectrum, centred at $(\frac{S_t}{2}, \frac{S_v}{2})$, as shown in figure 3(b). Signal B at frequency (T_B, V_B) , shown in figure 3(c), is a single vertical-temporal sinusoid at the same frequency as A 's alias due to interlacing, hence $T_B = (\frac{S_t}{2} - T_A)$ and $V_B = (\frac{S_v}{2} - V_A)$. Figure 3(d) shows the spectrum of B after interlacing. Comparison of figures 3(b) and 3(d) reveals that the magnitude spectra of A and B after interlacing are in fact the same. We synthesise these interlaced spectra with a progressive signal, X , which is the sum of two vertical-temporal sinusoids at the frequencies of A and B .

We wish to know, for all vertical and temporal frequencies, whether A or B is dominant in the interlaced version, i.e. whether humans perceive the pseudo-interlaced spectrum X to be derived from A or from B . If A is found to be dominant, we can deduce that an anti-alias filter should

pass A and suppress B . The result is that any signal components at A are interpreted correctly, and any components at B are removed instead of being misinterpreted. Using a set of A - B pairs that spans the spectrum, we obtain a compromise in the vertical and temporal resolution of the final interlaced video that is matched to the human response.

For our experiment we generated a set of test signals, each of which is a greyscale sinusoidal vertical grating moving upwards at a constant speed, so each signal is represented by a single point in the spectrum. For each pair of signals A and B , at frequencies (T_A, V_A) and $(\frac{S_t}{2} - T_A, \frac{S_v}{2} - V_A)$ respectively, we generated a video test clip at 50 frames per second which placed signals A and B next to $X=A+B$, as illustrated in figure 4. The final filter is intended to be applied after gamma correction, so the test signals are not further modified. We also do not correct for the display response, since the display is an important part of the reconstruction filter.

Subjects were asked to judge which of signals A and B is more similar to X , for each A - B pair. Additionally, the time taken for subjects to make a decision was recorded, in order to obtain a confidence measure.

3.1 Experiment Details

The set of test signals was sampled on a frequency grid that included 18 equally-spaced temporal frequency points spanning from 0 to $0.5 S_t$, and 16 equally-spaced vertical spatial frequency points spanning from $0.0294 S_v$ to $0.4706 S_v$. This results in equal frequency spacing in both dimensions relative to their respective sampling frequencies, and produces a set of 144 comparisons that span the vertical-temporal spectrum. Signals at zero spatial frequency and their aliases were not used since they could present a risk of seizures in susceptible individuals [14]. For the same reason, the test patterns were restricted to an area of 390×146 pixels. Each test signal had a duration of 10 s, and subjects were able to replay the signal if required. Subjects' first answers were recorded, and they were not able to change their decisions.

The test sessions included a short training phase of five comparisons for the subjects to acquaint themselves with the test interface and the kinds of signals that were to be used. The training samples were randomly drawn from the 144 test clips, with the restriction that one example of the lowest and one of the highest spatial frequency was always included. Six "dummy" tests, also drawn at random from the 144 real tests, were added to the beginning of the test phase, the answers for which were not recorded. Subjects therefore made a total of 155 comparisons during the test session.

The display and lighting conditions were calibrated according to recommendations for subjective assessments in a laboratory environment in ITU-R Recommendation BT.500-12 [15]. The display was a 32" Sony PVM-L3200 LCD monitor, and the tests were conducted at viewing distances of both three times the screen height (3H) and six times the screen height (6H), or 1.18 m and 2.35 m respectively. Subjects completed the tests at 3H and 6H in two separate sessions, with a minimum of 2 hours and a maximum of 3 days between them. There were 26 test subjects, all drawn from BBC Research and Development, 14 of whom are concerned with video picture quality as part of their normal work. Of the 26 subjects, 24 completed the test at both distances, one completed the test only at 6H, and one only at 3H, giving 25 sets of results for each distance.

3.2 Data Processing

Each set of results, i.e. the data from each subject-distance combination, consists of two functions of vertical-temporal frequency: a binary indication of whether or not each frequency was considered more similar to X than its alias, and the time taken to make the decision. The rating values for each subject were assigned numerical values: 1 for the dominant spectral position of each A - B pair, and -1 for its alias.

The time was recorded between subjects pressing 'play' and making a decision between A and B , regardless of whether they repeated the test clip after 10 s. In practice only 0.83% of times at

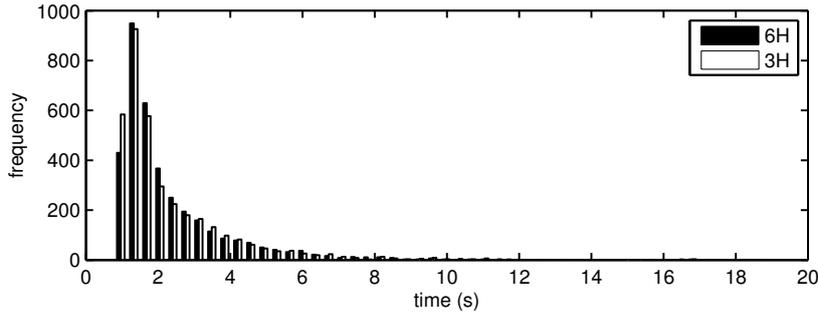


Figure 5: Histogram of raw response times for all subjects.

3H and 0.67% of times at 6H were over 10s (see figure 5), so due to the discontinuity in viewing for these very rare cases, any longer times were clipped to exactly 10s. Some subjects took longer than others on average to make decisions. In order to avoid bias towards any particular subject, we normalised the times for each subject-distance combination. All times in a set were divided by the median time for that set. This equalises the median time to 1 for all sets of results, whilst remaining insensitive to a small number of outliers, which could be caused by a subject being momentarily distracted.

The ratings were then weighted by dividing each rating value by the time taken to make the decision, thereby giving less influence to the decisions that took a long time. In fact the final binary filter templates are not significantly affected by the time-weighting procedure, but the confidence information will be valuable as error weighting for the next stage: calculating filter coefficients.

The number of test clips that could be presented was limited by potential viewer fatigue, so in order to obtain values sampled on a finer frequency grid for the filter specification process we applied cubic interpolation to the results, obtaining 8 times as many sampling points in both dimensions. The interpolated outputs were then smoothed using a Gaussian window in both dimensions. The window included 6 standard deviations of the Gaussian function and had length 33.

We then calculated means of the raw ratings (figures 6(a), 6(b)), normalised decision times (figures 6(c), 6(d)) and weighted ratings (figures 6(e), 6(f)) across all subjects.

3.3 Region-Growing Procedure for Filter Specification

The time-weighted ratings take the form of a continuous two-dimensional function of vertical-temporal frequency, as shown in figures 6(e) and 6(f). The final interlacing filter template should include the half of the vertical-temporal spectrum that was considered most dominant, i.e. the half with the highest rating. However, we have the additional restriction that the passband must fill a unit cell of the interlaced reciprocal sampling lattice, which means that the passband and stopband must be symmetrical. Although our results are intrinsically symmetrical, we wish to develop a method that would be suitable for alternative types of rating data, and in the general case simply choosing the top 50% of ratings would not guarantee symmetry. Instead we treat the spectral samples as pixels and develop an iterative region-growing procedure that maintains passband-stopband symmetry at every iteration.

Let $R(i, j)$ be the processed user ratings, with i and j indices along the temporal and vertical axes respectively. We begin by adding a single pixel to the filter passband, $R(0, 0)$, the point closest to the origin. Its alias, $R(M, N)$, where M and N are the total number of temporal and vertical frequency samples respectively, is added to the filter stopband, and all other pixels remain unassigned. For all unassigned pixels a cost of including them in the passband $C(i, j)$ is calculated, such that

$$C(i, j) = R(M - i, N - j) - R(i, j)$$

i.e. the cost of including pixel (i, j) in the passband is dependent on both its contribution to the passband, $R(i, j)$, and the contribution of its alias to the stopband, $R(M - i, N - j)$.

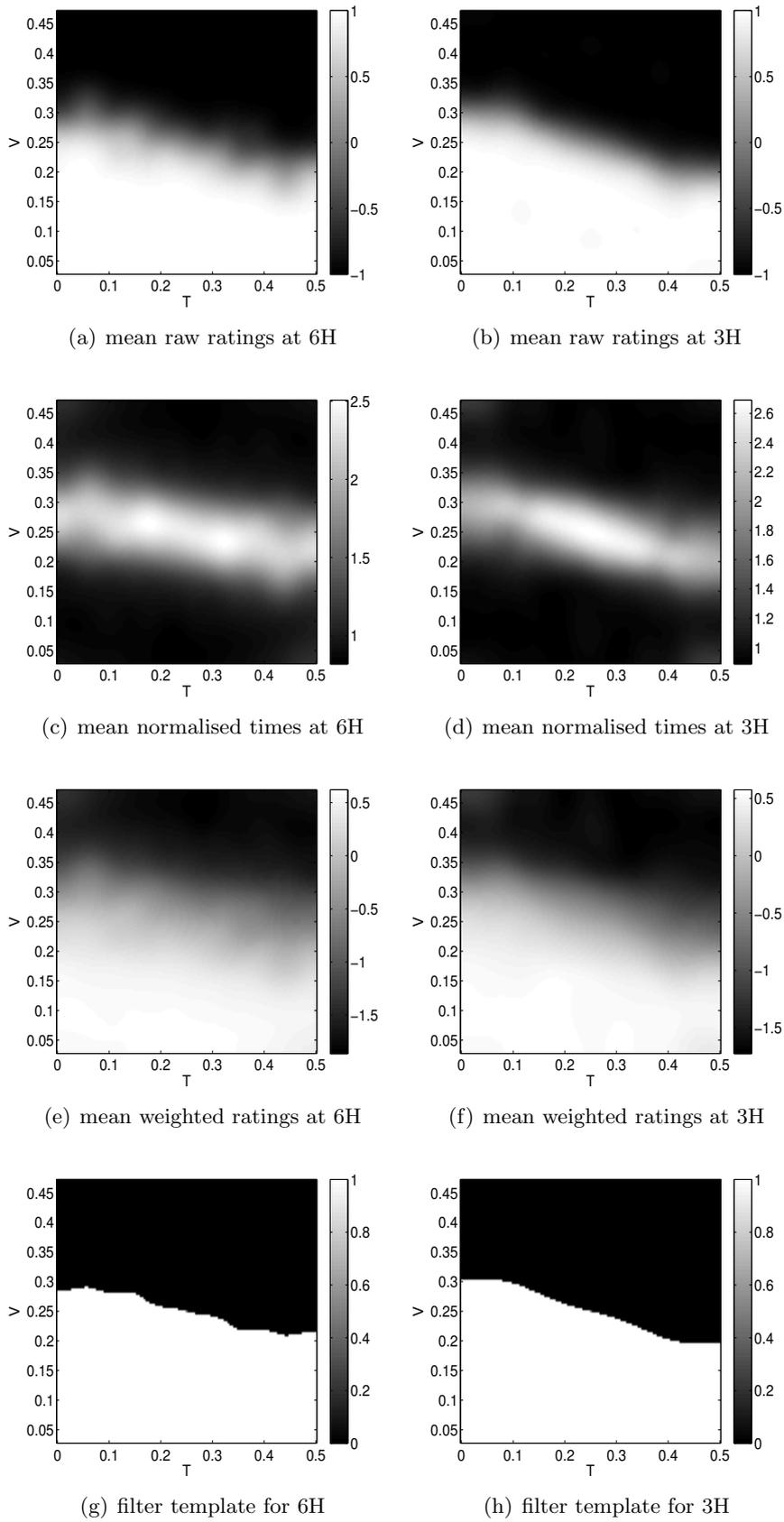


Figure 6: Results after interpolation and smoothing. Frequencies shown are relative to the respective sampling frequencies.

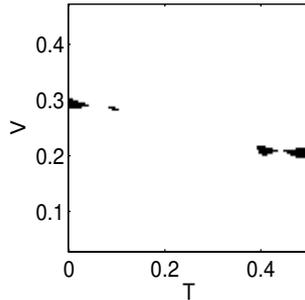


Figure 7: Intersection (shown in black) of the regions of significantly different weighted responses at 3H and 6H, and differences between the filter templates.

The region-growing procedure then iterates by comparing the cost of including each pixel that lies adjacent to the filter passband either vertically or horizontally (not diagonally), and includes the one with the lowest cost, removing it from the pool of unclassified pixels. At each iteration the alias of the most recently added pixel is committed to the stopband of the filter, and also removed from the pool. The process terminates when all pixels are committed to either the passband or the stopband, resulting in a binary image that can be treated as a filter design template.

4 Results

Figure 6 shows our results after processing. The average ratings across subjects are presented in figures 6(a) and 6(b). There is extremely strong agreement between subjects, with mean values extending to the extremes of 1 and -1, which shows us that our measurement method is effective. The only grey region in the mean ratings, indicating disagreement between subjects, is a narrow band around the potential boundary. This can be interpreted as a band of uncertainty, where it is not clear which of the two single sinusoids was dominant. Figures 6(c) and 6(d) show the mean decision times after normalisation, which each include a band of long decision times that corresponds to the band of disagreement in the raw ratings. This provides further support for uncertainty around the boundary, which means some flexibility could be permitted in the final filter magnitude response in this region.

Figures 6(e) and 6(f) show the time-weighted ratings, which resemble the mean ratings but with a more gradual transition over the boundary. The time-weighted ratings are used to derive the final filter templates shown in figures 6(g) and 6(h). The filter pass-bands have a common form in which low spatial frequencies tend to dominate, with a divergence from this tendency around $0.25 S_V$ where temporal frequency becomes more important. This means that we are prepared to sacrifice some spatial detail in moving objects in exchange for enhanced detail when the picture is stationary. The result is a boundary similar to that in figure 2(d), the only form shown in figure 2 that has taken perception into account.

The ratings at 6H and 3H are very similar. At 3H the filter passband extends slightly further up the spatial frequency axis at zero temporal frequency than at 6H, but the difference lies within the band of uncertainty indicated by the mean decision times at both 3H and 6H. Figure 7 shows the intersection of the regions where the filter templates disagree, and the regions where the weighted ratings are significantly different according to a matched t-test ($p < 0.01$) on the 24 sets of ratings that come from the same subjects at both distances. This area is very small, and almost certainly smaller than the resolution achievable by filters of moderate length, so we conclude that the same filter would be suitable for both viewing distances: a beneficial result for real broadcast systems where the viewing distance cannot be controlled.

5 Conclusions

We have presented a novel test paradigm for measuring the relative sensitivity of the human visual system to pairs of vertical-temporal frequencies. The paradigm is simple and powerful, as shown by the extremely strong agreement between test subjects. We have also implemented a new region-growing technique to derive an interlacing filter template from the collected data, which is specifically tailored to the task of meeting the interlacing filter constraints of filling a unit cell of the interlaced reciprocal sampling lattice.

We conducted an experiment according to the new paradigm, and hence derived an optimal interlacing filter template. In the template low spatial frequencies tend to be dominant regardless of the temporal frequency, but there is a region around a quarter of the original vertical sampling rate in which the temporal frequency dominates. This frequency response approximates that of a filter presented by Weston as a possibility for a de-interlacer [9] on heuristic grounds, though with slightly less curvature, so we have therefore provided some scientific grounding for Weston's de-interlacing filter as well as measuring the optimal magnitude response. We conducted the experiment at two viewing distances, but the resulting filter templates are not significantly different, so we conclude that the target viewing distance is not critical to the filter design.

The current work only addressed our sensitivity to luminance, and it remains to determine whether the same filter would be most suitable for interlacing chrominance signals. However, it is well-known that we are less sensitive to colour detail than brightness [16], so it is likely that the precise filter response is less critical for chrominance. Additionally, in order to avoid separating luminance and chrominance for objects that happen to move at a particular speed, the same filter should be applied to all three components. Therefore at this stage we would recommend applying these results to chrominance as well as luminance.

The filter templates can be regarded as optimal in the sense that a filter exactly matching the template would prevent overlapping aliases while passing the more visible of each baseband-alias frequency pair. However, it is possible that some aliasing is in fact preferred to softening of the image in the context of real video. Hence, our filter templates can be regarded as a principled starting point for designing a family of possible interlacing filters. In order to fully evaluate their suitability, it will be necessary to design real filters from the templates, and measure the perceived visual quality of a range of video material after interlacing using our filters and existing ones. These experiments are the subject of our further work.

6 Acknowledgements

The author would like to thank Richard Salmon and Alastair Bruce for their assistance with calibrating the display and lighting for the experiments.

References

- [1] C. K. P. Clarke and N. E. Tanton, "Digital standards conversion: Interpolation theory and aperture synthesis," BBC Research Department Report RD 1984/20, Dec. 1984.
- [2] E. Dubois, "The sampling and reconstruction of time-varying imagery with application in video systems," *Proc. IEEE*, vol. 73, no. 4, pp. 502–522, Apr. 1985.
- [3] R. C. Ballard, "Television system," US patent 2,152,234, 1932.
- [4] R. D. Kell, A. V. Bedford, and M. A. Trainer, "Scanning sequence and repetition rate of television images," *Proc. Institute of Radio Engineers*, vol. 24, no. 4, pp. 559–576, Apr. 1936.

- [5] T. Cooklev, T. Yoshida, and A. Nishihara, “Maximally flat half-band diamond-shaped FIR filters using the Bernstein polynomial,” *IEEE Trans. Circuits and Systems II: Analog and Digital Signal Processing*, vol. 40, no. 11, pp. 749–751, Nov. 1993.
- [6] G. de Haan and E. B. Bellers, “Deinterlacing—an overview,” *Proceedings of the IEEE*, vol. 86, no. 9, pp. 1839–1857, Sep. 1998.
- [7] T. Borer, “Television standards conversion,” Ph.D. dissertation, University of Surrey, Oct. 1992.
- [8] A. Roberts, *Circles of Confusion*, 1st ed. EBU, 2009.
- [9] M. Weston, “Interpolating lines of video signals,” US Patent 4,789,893, Dec. 1988.
- [10] M. Weston and D. M. Ackroyd, “Fixed, adaptive, and motion compensated interpolation of interlaced TV pictures,” in *Proc. International Broadcasting Convention (IBC)*, Sep. 1988, pp. 220–223.
- [11] S. H. Keller, F. Lauze, and M. Nielsen, “Deinterlacing using variational methods,” *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2015–2028, Nov. 2008.
- [12] J. G. Robson, “Spatial and temporal contrast-sensitivity functions of the visual system,” *J. Opt. Soc. Am.*, vol. 56, no. 8, pp. 1141–1142, 1966.
- [13] D. H. Kelly, “Motion and vision. II. Stabilized spatio-temporal threshold surface,” *J. Opt. Soc. Am.*, vol. 69, no. 10, Oct. 1979.
- [14] C. D. Binnie, J. Emmett, P. Gardiner, G. F. E. Harding, D. Harrison, and A. J. Wilkins, “Characterising the flashing television images that precipitate seizures,” *SMPTE Journal*, pp. 323–329, Jul./Aug. 2002.
- [15] ITU-R, “Methodology for the subjective assessment of the quality of television pictures,” Recommendation ITU-R BT.500-12, 2009.
- [16] W. Pratt, “Spatial transform coding of color images,” *IEEE Trans. Communication Technology*, vol. 19, no. 6, pp. 980–992, Dec. 1971.