UNIVERSITY OF BRISTOL

University of Bristol - Explore Bristol Research
General rights

Contrast Sensitivity of the Wavelet, Dual Tree Complex Wavelet, Curvelet, and Steerable Pyramid Transforms

Paul Hill, *Member, IEEE*, Alin Achim, *Senior Member, IEEE*, Mohammed E. Al-Mualla, *Senior Member, IEEE*, and David Bull, *Fellow, IEEE*

Abstract—Accurate estimation of the contrast sensitivity of the human visual system is crucial for perceptually based image processing in applications such as compression, fusion and denoising. Conventional contrast sensitivity functions (CSFs) have been obtained using fixed-sized Gabor functions. However, the basis functions of multiresolution decompositions such as wavelets often resemble Gabor functions but are of variable size and shape. Therefore to use the conventional CSFs in such cases is not appropriate. We have therefore conducted a set of psychophysical tests in order to obtain the CSF for a range of multiresolution transforms: the discrete wavelet transform, the steerable pyramid, the dual-tree complex wavelet transform, and the curvelet transform. These measures were obtained using contrast variation of each transforms' basis functions in a 2AFC experiment combined with an adapted version of the QUEST psychometric function method. The results enable future image processing applications that exploit these transforms such as signal fusion, superresolution processing, denoising and motion estimation, to be perceptually optimized in a principled fashion. The results are compared with an existing vision model (HDR-VDP2) and are used to show quantitative improvements within a denoising application compared with using conventional CSF values.

Index Terms—Discrete wavelet transforms, contrast sensitivity function.

I. INTRODUCTION

THE Contrast Sensitivity Function (CSF) measures the threshold sensitivity of simple grating patterns of the Human Visual System (HVS) over a wide range of spatial frequencies. It plays a central role in the visual models and metrics used in perceptually-based image processing techniques such as image fusion and compression (e.g. [1], [2]).

Wavelet decompositions (see section II) provide invertible multiscale transforms enabling two-dimensional

spatial-frequency analysis. The usage of wavelet decompositions in perceptual image processing has commonly been based on standard CSFs that have been generated using fixed-size Gabor functions (e.g. [2]). Watson et al. [3] measure and model the CSF curves associated with the DWT and observe that they do not follow the conventional CSF shape. Specifically, maximum contrast sensitivity is significantly shifted toward lower frequencies.

Although effective in multiscale decomposition, the basis functions of the DWT do not effectively decompose orientated two dimensional content compared to more recent wavelet decompositions (as discussed below). Our work is based on the work of Watson et al. [3] and its aim is therefore to generalise this previous work; specifically, to generate contrast sensitivity measures for each subband for a range of very effective and popular wavelet transforms, thus enabling more effective and accurate perceptual processing. Additionally, we have added further appropriate and useful experimental conditions (such as pink noise) compared to Watson et al. [3]. In principle, the presented thresholds could be predicted using Linear System Theory (LST) from previous developed models. However, due to the inherent non-linearities of the HVS and related effects, direct measurements will provide precision to an accuracy not possible through model predictions.

This paper is structured as follows. First the background of both wavelets and contrast sensitivity functions are discussed in section II. The methods used to obtain the threshold values for each of the considered transforms is described in section III. The results are presented and discussed in section IV. This is followed by a section comparing the results to an existing vision model in section V and a description of applications in section VI. Finally a conclusion including a description of further work is included in section VII.

II. BACKGROUND

A. Wavelet Transforms

Wavelets, in their different forms, provide a large number of invertible transforms that decompose a signal into a self-similar, weighted set of basis functions that vary in scale and orientation (in two dimensions). Although the discrete form (the Discrete Wavelet Transform: DWT) has primarily been used for compression, more recent non-critically decimated discrete forms (i.e. where the number of coefficients equals the number of samples) have been defined for the

use in analysis applications such as: image and signal fusion, super-resolution processing, denoising and motion estimation. Popular wavelet transforms used for analysis include the DWT, DT-CWT, Curvelet and Steerable Pyramid Transforms. These are discussed below.

1) Discrete Wavelet Transform (DWT): The DWT is a critically decimated invertible transform that uses separable orthogonal or bi-orthogonal high and low pass spatial filters that have been designed to provide perfect reconstruction [4].[1] The CSF has been measured and modelled for the DWT by Watson et al. [3].

2) Dual-Tree Complex Wavelet Transform (DT-CWT): The two dimensional DT-CWT was formulated by Kingsbury [5] and Selesnick *et al.* [6] in order to provide improved directional resolution and shift invariance compared to the conventional DWT. In common with the DWT, the DT-CWT is implemented using separable FIR filters over octave frequencies using subsampling between levels. However, the DT-CWT utilises two separate trees that form Hilbert filter pairs leading to complex real and imaginary coefficients in each subband.[2] The transformation is able to distinguish between positive and negative frequencies leading to 6 orientated complex subbands at each scale (with the 6 orientated subbands being evenly angularly distributed as illustrated in figure 3).

It has been identified by Adelson et al. [9] that separating the positive and negative frequencies of an image transform is important in terms of human vision based processing. Figure 3(a) shows 3 subbands' real basis functions (there are 6 per scale with the remaining 3 being mirrors of the ones shown). This figure shows: the basis functions orientated at the 3 angles $(15°, 45°, 75°)$, exponential change in support and how the basis functions resemble Gabor functions. The near shift invariance and improved directional selectivity have facilitated excellent results in denoising, fusion and other image processing applications [10]–[12].

3) Curvelet Transform: Conventional wavelet transforms such as the DWT are not well suited to two dimensional discontinuities. The Curvelet transform is a multiscale decomposition where an image is decomposed with geometric bases in multiple directions and positions [13]. Specifically, frequency space is covered with directionally selective subbands at octave frequency centres with increasing directional selectivity as the frequency increases (leading to needle shaped basis functions at these finer scales).

Curvelets use parabolic scaling whereby at scale j each basis function has approximately the length of $2^{-j/2}$ and width 2^{-j}. This arrangement leads to geometrically tuned basis functions that are better able to represent orientated discontinuities (especially at higher frequencies) and therefore give a sparser representation for natural images. Due to their geometric nature they have given very good results in

applications such as denoising, image contrast enhancement and fusion (e.g. [14], [15]). Curvelets have been implemented using both an Unequally-Spaced Fourier Transform (USFFT) and also the wrapping of specially selected Fourier samples (wrapping method) [13]. The USFFT implementation has been used within this work (to generate the basis functions shown in figure 5(a)).

4) Steerable Pyramid Transform: The Steerable Pyramid decomposition is a multi-orientation, multi-scale and invertible transform that has been used for many image processing applications over the last two decades [16], [17]. The Steerable Pyramid differs from the DWT, DT-CWT and Curvelet transforms in that the degree of orientation selectivity can be arbitrarily selected. Using kth order directional derivatives, the basis functions have $k+1$ orientations over an arbitrary number of scales and the transform was originally implemented using non-separable 2D steerable spatial filters. We have used the FFT implementation [17] to generate the basis functions shown in figure 4(a).

B. Contrast Sensitivity Function

A widely adopted approach to obtain the transfer function of the HVS is to measure the contrast sensitivity for spatially modulated and orientated Gabor or sinusoid gratings. CSF functions obtained with this method produce the typical inverse u-shaped graph that illustrates the HVS attenuation in sensitivity at high and low frequencies. A large number of models for the CSF have been developed such as those by: Daly [18], Barten [19] and van Meeteren and Vos [20].[3]

These works use common methods for obtaining the CSF using sinusoidal or Gabor gratings and vary the pattern contrast on a fixed uniform background luminance to obtain the threshold when the pattern is "just seen". Experiments commonly use iterated forced choice methods such as 2AFC [3] or 4AFC [21] to obtain this threshold. The current contrast together with whether the observer has chosen the correct forced choice is then compared to a chosen psychometric function model producing an updated contrast value and the method iterated until convergence. It should be noted that the estimated psychometric function represents the probability of detection reflecting that HVS thresholds are not simple step functions.

1) Contrast Sensitivity Function Dependencies: The CSF has a large number of dependencies on signal context. We now give a non exhaustive list of these dependencies.

a) Static average background luminance: Many different studies have measured the CSF on various (uniform) luminance backgrounds (for example an extensive study of CSF for different backgrounds on HDR displays was conducted by Kim et al. [21]). However, for non HDR applications, a static average background luminance is often assumed. Within the following experiments the background was measured to be $23.91 cd/m^2$. This value was chosen to give a suitable average background luminance for a non HDR display.

---

[1]The relationship between orthogonality and critical / non-critical decimation is complicated within wavelet transforms. Specifically, each pair of filters can either be orthogonal or non-orthogonal. When using non-critical decimation, the transform as a whole is not orthogonal but this may be from multiple effects such as the type of filters selected and the filter structure.

[2]Previous work by Watson [7] has considered complex coefficients and the role of the Hilbert Transform for multiscale vision-motivated transforms (the cortex transform). This work has recently been extended by Lukin [8].

[3]These three references have all used sinusoids windowed with a rectangle. This is historically the first way that CSFs were obtained before the more common contemporary use of Gabor filters.

b) Adaptation effects: The effect of light adaptation on the CSF was first investigated by Van Nes *et al.* [22]. More recent contrast sensitivity experiments have used a selection of typical natural images to adapt the HVS to typical frequency and spatial content. In studies, these images have been presented before, [23] and during, the stimulus presentation [23], [24]. A study of the effect on the CSF functions on natural image content is investigated by Johnson and Fairchild [25]. Many other studies have shown the adaptation of the varying stimuli to other controlled (orientated or scaled) stimuli such as masking Gabor functions (e.g. [26]). A more conventional approach (single stimulus presentation) has been used in the work presented here as it was considered that the analysis of each basis function in any application will be made alone without any reference to global frequency. Further independent masking effects could be used to extend the model in future work. However, simple $1/f$ spatial adaptation, using pink noise (see section III-E), was employed to give some basic adaptation in a non-localised sense.

c) Suprathreshold effects: For applications such as compression and fusion, quantitative measures of difference (for contrast sensitivity) above the threshold of perception are important. This has rarely been dealt with in the literature. Most contrast sensitivity models for compression assume a linear response above the threshold. However, it has been demonstrated that the effects at threshold are significantly different than at suprathreshold levels [27], [28] leading to the conventional CSF curves not being applicable at levels far from threshold. Specifically, Daly [29] and Georgeson and Sullivan [27] have shown that equal contrast curves flatten out as the contrast becomes significantly greater than the threshold level. This effect is termed contrast constancy.

d) Orientation: It has been observed that natural scene imagery contains a larger proportion of frequency content in vertical and horizontal orientations (i.e. the cardinal axes) relative to the oblique orientations [30], [31] (with an additional slight bias to horizontal content as noted by Hansen and Essock [32]). This is reflected in the orientational tuning of the contrast sensitivity function of the HVS. Initial studies have shown that the contrast sensitivity function obtained using simple grating thresholds is more sensitive to horizontal and vertical orientations (the so called "Oblique Effect" [33], [34]). However, more recent experiments [30], [32] using broadband natural image masking content have conversely shown that the visual system is more sensitive to oblique orientations. It is anticipated that the conventional "Oblique Effect" will be reflected in our results. Therefore threshold data has been included for different basis function orientations where practical.

e) Additional dependencies: These include phase, chrominance, and temporal effects. These effects are not further investigated in our work.

Due to our application focus and to ensure tractability of analysis, we assume only a simple dependence of the CSF on orientation and $1/f$ spatial adaptation. Further work is intended to extend the model to take into account the remaining dependencies listed above.

III. METHODS

A. Display and Viewing Conditions

1) Display Visual Resolution: The thresholds of the different transform basis functions depend upon both the viewing distance and the display visual resolution in pixels/degree. Given a viewing distance v in cm and a display resolution d in pixels/cm, the display visual resolution r in pixels/degree is given by (1).

$$r = dv \, \tan\left(\frac{\pi}{180}\right) \approx dv\frac{\pi}{180} \approx \frac{dv}{57.3} \qquad (1)$$

For each trial, the viewing distance was set to 57cm therefore giving $r \approx d$. Given this viewing distance and the resolution of the monitor, the display visual resolution was 21.38 pixels/degree. This was used for all subsequent calculations. This resolution relates to the design specification for HDTV (approximately 30 cycles/degree). It should be noted that there are significant differences in sensitivity as a function of the viewing distance independent of the visual resolution in pixels/degree. The distance used within our experiments would therefore be more applicable to desktop use than cinema or large TV applications.

B. Basis Function Spatial Frequency

In order to obtain meaningful CSF curves, the spatial frequency (cycles/degree) can be easily calculated for conventional Gabor gratings, since their central frequencies can be specifically defined. For the basis functions used throughout our work, the spatial frequency is calculated by (2).

$$f = \frac{r}{c} \qquad (2)$$

where f is the spatial frequency in cycles/degree, r is the display resolution in pixels/degree and c is the cycles/pixel associated with the basis function. In previous work by Watson et al. [3], c is defined as 2^λ where λ is the wavelet level. This definition is not applicable to our work because each subband and its associated basis function will have a variable central frequency that changes according to both scale and orientation. The central frequency c for our work is calculated for each subband as the basis functions' maximum spatial frequency (as a proportion of the sampling frequency). This was calculated by finding the distance (in normalised units of frequency) from DC to the maximum magnitude of the two dimensional frequency transform of the subbands' basis function.

C. Display Calibration

In order to generate accurate contrast sensitivity threshold values, the monitor output must be calibrated so the stimuli provide linear luminance differences and contrast. The chosen monitor (see below) was calibrated using a photometer and the calibration functions provided within Psychtoolbox [35]. The gamma value was measured to be approximately 1.9 (generated from a range of 9 luminance values across the monitor's possible values). A look-up table was used in all

experiments thus creating a linearised display. A linearised display was chosen due to the following reasons:

- As illustrated by Watson et al. [3] the difference in thresholds for a linearised and non-linearised display is approximately an offset. As the generated models were monotonically increasing, the relative offsets between transform subbands are not considered to be significant for comparative analysis.
- For compatibility with compression / analysis methods that have assumed linearised displays and where visual models are combined under the assumption of linearity (e.g. Höntsch and Karam [36]).
- Potential variations in gamma within display applications (typical values from 1.6 to 2.6).

Our results should only strictly be valid for linearised displays although conversion from linearised to non-linearised threshold models is possible using the methods given by Watson et al. [3] and Peterson et al. [37].

D. Stimuli

Basis functions from the DWT, DT-CWT, Steerable Pyramid and Curvelet transforms were used as the stimuli for the threshold measurement experiments described in section III-E. For all transforms, the subbands are referenced as $\{\lambda, \theta\}$ where λ is the decomposition level and θ is the subband orientation. For transforms using complex basis functions, the real component was used in all cases.

For all transforms, the spatial basis function images were created as follows.

- A 512×512 zero filled image was transformed using the chosen transform (resulting in all zero coefficients within each subband).[4]
- A single centrally located coefficient in the chosen subband was set to 1 (or 1+0i in the complex transform case).
- The transform was inverted and the basis function exactly centred to produce the basis function stimulus.

The basis function stimuli were presented in a manner identical to the work of Watson et al. [3]. Firstly, the basis functions were normalised to give unit maximum amplitude (resulting in a signal in the range of $[-1,1]$). These normalised basis functions were variably scaled to produce maximum amplitudes in the range of $[0,128]$. The scaling factor of these signals is our measure of stimulus contrast ΔY. Before presentation (within the threshold calculation procedure described in section III-E) the scaled signal is added to the display average value of 128 and the resulting image in the range $[0,255]$ is presented on the linearly compensated display. Our defined stimulus threshold ΔY is therefore the maximum amplitude of the signal minus 128 at the threshold of perception.

The average intensity of the display L_0 (when displaying a uniform value of 128) was found to be $23.91 cd/m^2$ and the maximum intensity of the display (displaying a uniform value of 255) was found to be double this value (due to the linearised display). The commonly used contrast measure

---

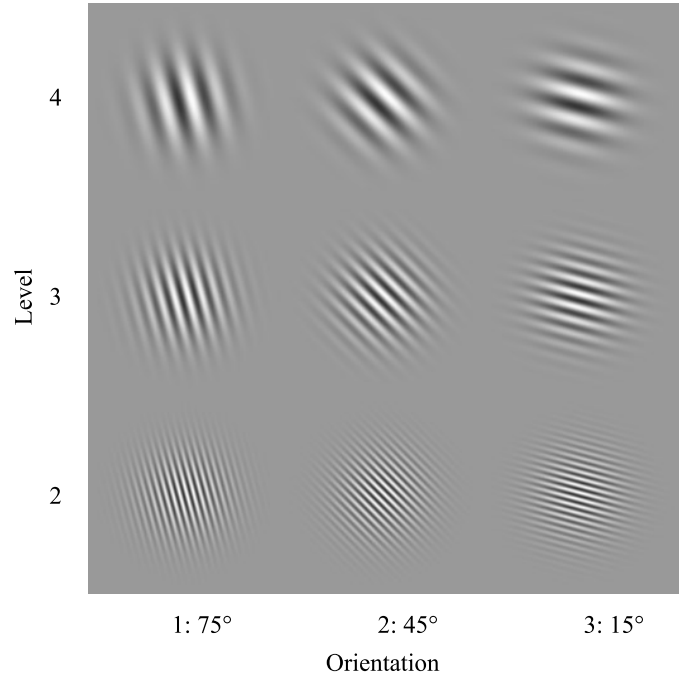[4]This size was found to be large enough for all transforms' basis function to fit within for up to 5 scales.

Fig. 1. Gabor Function Stimuli.

$\Delta L/L_0$ has been used (in dB units) defined for threshold experiments by Manituk et al. [38] as $C = 20log_{10}\left(\frac{\Delta L}{L_0}\right)$. For our experiments, due to the linearised display, ΔL is related to ΔY as:

$$\Delta L = L_0 \Delta Y / 128 \qquad (3)$$

The stimulus contrast ΔY defined above can therefore be converted to the contrast C using

$$C = 20log_{10}\left(\Delta Y / 128\right) \qquad (4)$$

All stimuli were displayed on a calibrated SONY Trintion 20" monitor (Model number 300SF).

1) DWT Basis Function Stimuli: A selection of basis DWT functions is shown in figure 2(a) ($\{2\ldots4, 1\ldots3\}$). The actual basis functions used in the experiments were $\{1\ldots5, 1\ldots3\}$ i.e. all three orientations for the first 5 levels. The filters used were the nearly symmetric filters designed by Abdelnour and Selesnick [39]. These filters were used as they were the default first level filters used in the DT-CWT in section III-D2.

2) DT-CWT Basis Function Stimuli: Figure 3(a) shows the basis functions labelled $\{2\ldots4, 1\ldots3\}$. This figure shows how the positive and negative frequencies can be differentiated by the DT-CWT, leading to 6 orientated subbands at each level ($+/-15°, +/-45°, +/-75°$). Figure 3(b) shows how these basis functions are related to their frequency response zones. The basis functions used in the experiments were $\{1\ldots5, 1\ldots3\}$ i.e. all the scales from 1 to 5 but only the first three orientations. A subset of orientations was used to make the experiments tractable. We have made the assumption that sensitivity to a stimulus is not affected by left-right reflection. Therefore, as the basis functions for the DT-CWT, Curvelet and Steerable Pyramid are reflected across the y-axis, the results are assumed to hold for the reflected orientations.
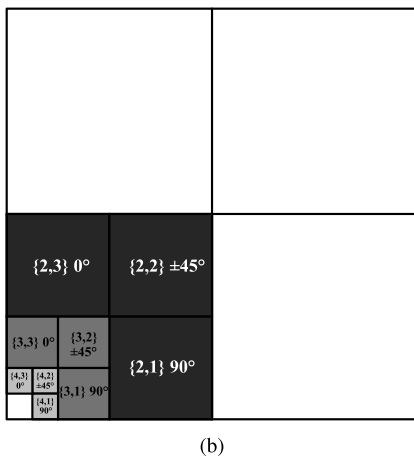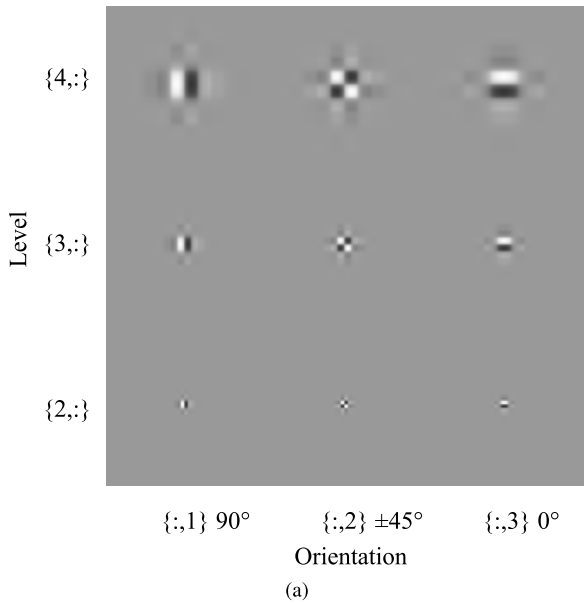
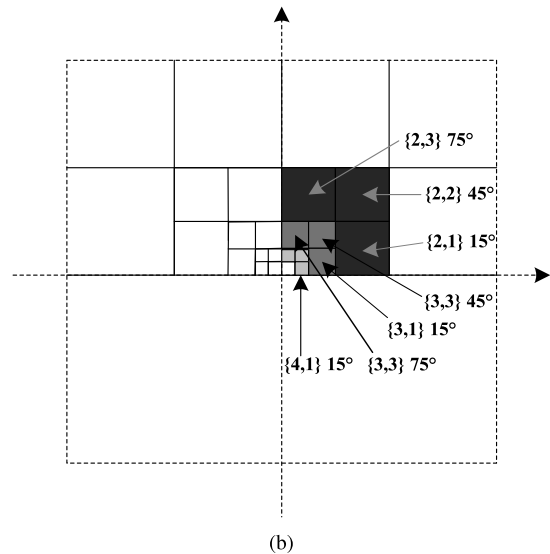Fig. 2. DWT Decomposition: Subband labels {level,orientation}. (a) DWT Basis Functions. (b) DWT Spatial Subbands.

Fig. 3. DT-CWT Decomposition: Subband labels {level,orientation}. (a) DT-CWT Basis Functions. (b) DT-CWT Subbands in the Frequency Domain.

3) Steerable Pyramid Basis Function Stimuli: The Steerable Pyramid basis functions and their frequency regions are shown in figures 4(a) and 4(b) respectively. The basis functions displayed are {2 . . . 5, 1}. The basis functions used for the experiments are {1 . . . 5, 1}. The Fourier domain implementation of the Steerable Pyramid transform was chosen (as it gives perfect reconstruction). The number of orientations of the transform can be chosen arbitrarily. However, we have used four orientations as this gives good directional coverage of the frequency domain and is the default implementation number within the available MATLAB code.[5]

4) Curvelet Basis Function Stimuli: The curvelet basis functions used in the experiments are shown in figure 5(a) and can be denoted {2 . . . 5, 1}. Only 4 scales were used as the single residual highpass filter did not contain any orientational information. Only one orientation per scale was selected for the experiment due to the very large number of separate orientations (16 orientation bands for scales 2 and 3 and 32 orientation bands for scales 4 and 5). Analysis of all these

orientations (even with assumed left-right mirroring) would make the experiment intractable. This is a different case from the Steerable Pyramid orientations as the Steerable Pyramid filters (at the same scale) are rotated versions of each other but the Curvelet filters are not. However, the differently orientated Curvelets at the same scale do resemble each other so the final CSF values for all directions are generalised from the single orientation but interpolated to the actual filter frequency centres of each subband. Threshold values are therefore available on a per-subband basis within the MATLAB code (discussed below).

5) Gabor Function Stimuli: The Gabor functions shown in figure 1 are included for reference and completeness. The centre frequencies (measured from DC) of the Gabor functions are equivalent to the centre frequencies of the horizontal (and vertical) subbands of the DWT (as assumed given by Watson et al. [3]). The spatial support of the Gabor Stimuli is defined by a constant Gaussian envelope with the sigma

---

[5]Steerable Pyramid code available at www.cns.nyu.edu/~eero/steerpyr

Fig. 4. Steerable Pyramid: Subband labels {level,orientation}. (a) Steerable Pyramid Basis Functions. (b) Steerable Pyramid Subbands in the Frequency Domain.
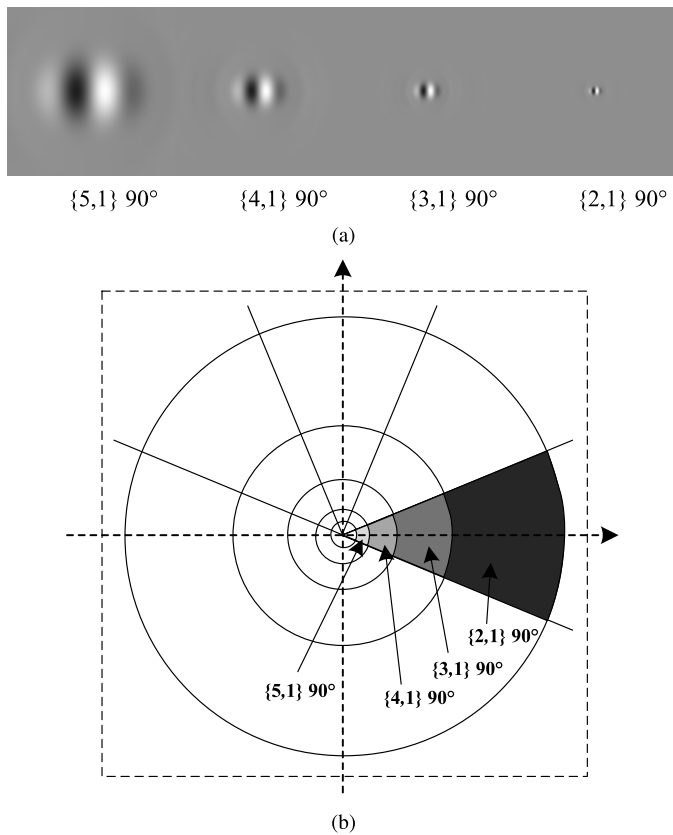

Fig. 5. Curvelet Transform: Subband labels {level,orientation}. (a) Curvelet Basis Functions. (b) Curvelet Subbands in the Frequency Domain.

equal to 1 visual degree (this compares to 0.5 visual degree for equivalent patterns within Modelfest [40] and 1.5 visual degrees for the Gabor patterns used by Kim et al. [21]).

E. Threshold Measurement

It has been recognised that forced choice experiments are an effective way to determine contrast sensitivity at threshold [21], [41]. Kim et al. [21] have used a 4AFC method to derive CSF curves, stating that the convergence to the perceptual threshold is faster than using 2AFC. However it was considered that presentation of stimuli at 4 possible spatial positions would confuse the spatial and temporal aspects of the results and therefore centrally located stimuli using the 2AFC method was chosen.

The 2AFC trials consisted of two 200ms intervals separated by 500ms [3]. One interval contained the stimulus added to a uniform grey screen and the other interval contained just the uniform grey screen (randomly alternated after each trial). Each interval was signalled by a separate audible warning tone. After each presentation, the observer was forced to select which interval they considered to contain the stimulus (using two mouse buttons).

Between presentation trials, the amplitude of the stimulus ΔY was varied adaptively using a Quest staircase [42]. At the end of the trials, for each stimulus a Weibull distribution was fitted to the proportion correct. The final threshold was estimated as the stimulus amplitude giving an 82% correct rate [43].

The code for our experiments used the Psychtoolbox [35] for MATLAB (version 3.0.11), utilising the toolbox's stimulus display, stimulus timing (with audio warnings) and QUEST implementations.

Other features of the trials include:
- 36 trials for each stimulus.
- The threshold was not adapted for the first 4 trials for each stimulus
  - This was done because: (i) observers became used to the experimental procedure and (ii) mistakes within the first few trials may lead the Quest method to erroneous threshold values that may be difficult to recover from.
- Between each interval and trial, a small cross (5 × 5 pixels) was presented at the centre of the monitor screen (luminance value 0). As the stimuli were also presented at the centre of the screen this "fixation point" was used to aid the observer in keeping the presentation within the observer's foveal region. The cross pattern was extinguished while the stimuli were presented.
- Between each trial an adaptation image was displayed. This consisted of achromatic spatial $1/f$ pink noise [44] with mean luminance equal to that of the stimuli (and background), a maximum amplitude of 128 and a standard deviation of 29.10.
  - This helped keep the observers focused on the screen between presentation of the stimulus [45].

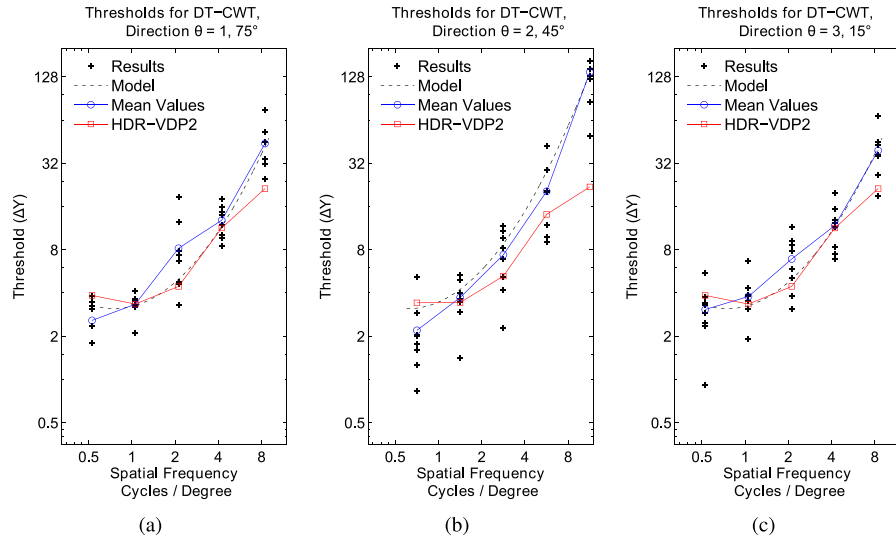Fig. 6. DT-CWT threshold results ΔY showing actual datapoints, means, model and HDR-VDP2 model results. (a) Subbands {1…5,1} 75°. (b) Subbands {1…5,2} 45°. (c) Subbands {1…5,3} 15°.
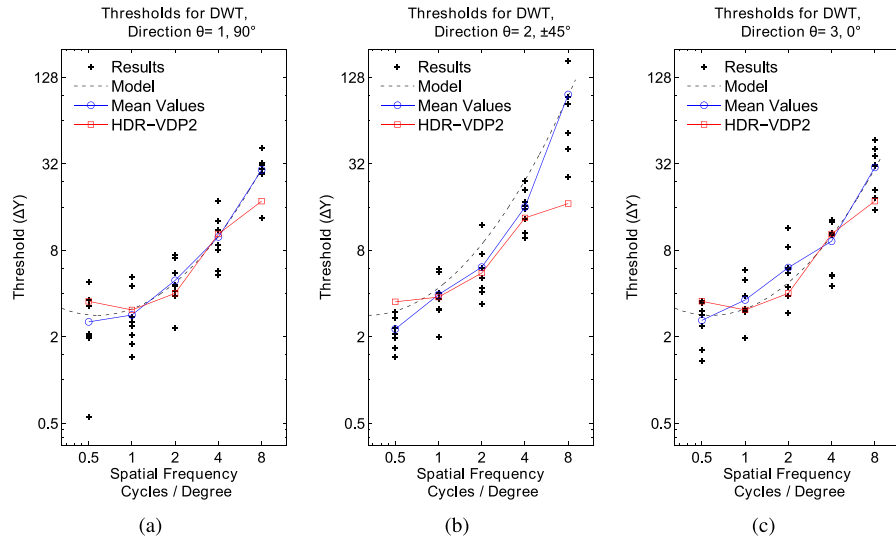
Fig. 7. DWT threshold results ΔY showing actual datapoints, means, model and HDR-VDP2 model results. (a) Subbands {1…5,1} 90°. (b) Subbands {1…5,2} ±45°. (c) Subbands {1…5,3} 0°.

- For each observer, before the trials the same adaptation image was displayed for 60 seconds in order to adapt the observers' visual system.
- A pink noise adaptation visual image was used to adapt the observers' visual system to the typical frequency content of natural images (in a statistical sense).
- There were eight observers. All observers were assessed to have normal or corrected to normal vision. None of the observers wore contact lenses. Viewing was within a darkened room and binocular.
- Causes spatial frequency adaptation.

IV. RESULTS

Figures 6, 7, 9 and 10 show the threshold values for the transform subband basis functions described above. Eight participants were used in these experiments and

significant outliers were discounted as a pre-processing stage. This was done by discounting samples outside confidence limit of 95% (i.e. an alpha of 0.05) as described in detail by Grubbs [46] and implemented using the `deleteoutliers.m` MATLAB function.[6] The confidence limits were calculated across all subjects for each stimulus. The crosses in each graph show the data after outlier removal. The average value over the eight observers is shown as a solid line on each graph. The graphs are plotted using log-log axes in order to better portray the variance over the range of frequencies and thresholds.

A. Threshold Model

Although there are a multitude of possible different parametric formulae available to model transform thresholds, we

[6]http://uk.mathworks.com/matlabcentral/fileexchange/3961-deleteoutliers/content/deleteoutliers.m
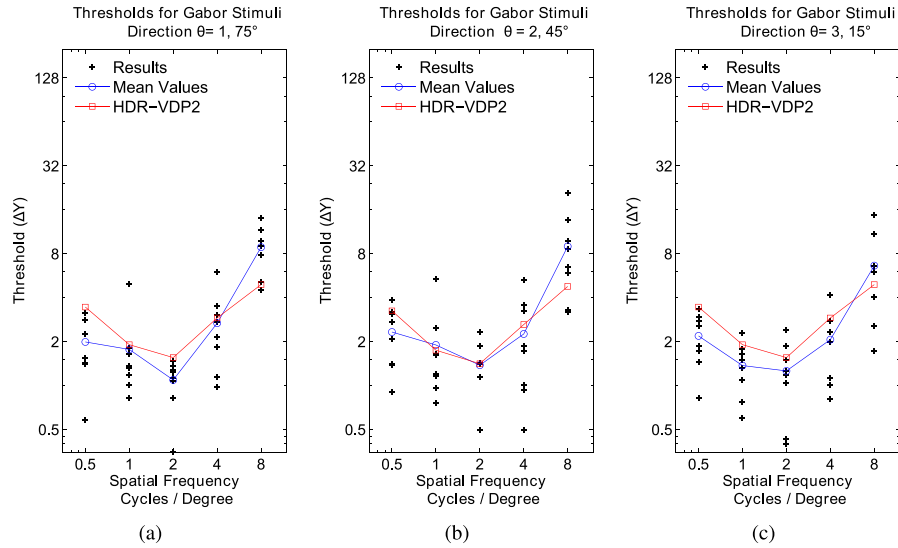
Fig. 8. Gabor threshold results ΔY showing actual datapoints, means and HDR-VDP2 model. (a) Subbands {1…5,1} 75°. (b) Subbands {1…5,2} 45°. (c) Subbands {1…5,3} 15°.
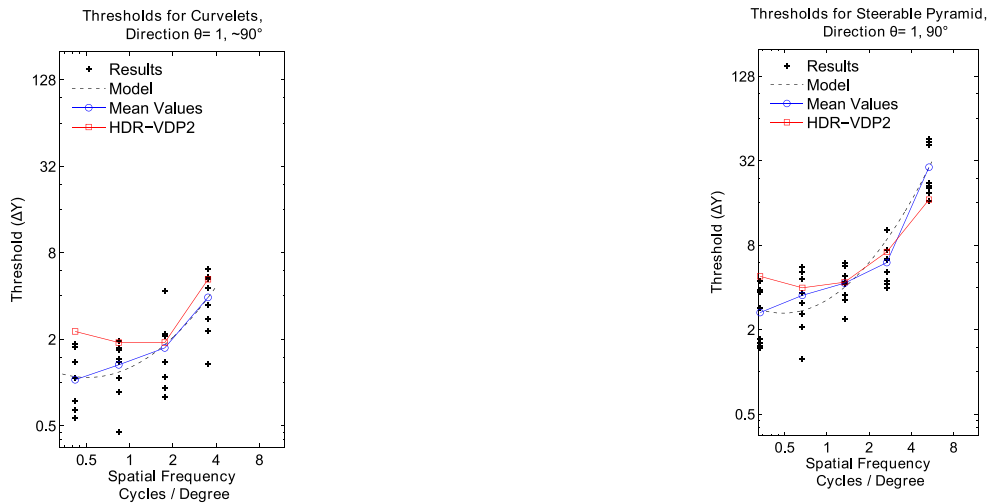
Fig. 9. Curvelet thresholds ΔY showing actual datapoints, means, model and HDR-VDP2 model results.

Fig. 10. Steerable Pyramid thresholds ΔY showing actual datapoints, means, model and HDR-VDP2 model results.

have adopted the model used by Watson et al. [3] shown in equation (5). This equation represents a parabolic curve in logarithmic space of the frequency and threshold.

$$\log\left(Y\right) = \log(a) + k\left(\log f - \log g_\theta f_0\right)^2. \qquad (5)$$

Where a is the contrast offset, g_θ is the orientation subband frequency offset, f_0 is the frequency minimum and k is the scaling product of the exponent.

The dotted line on each graph shows a fitted curve using standard least squares regression. The curves and average plots do not coincide for all the graphs as: firstly, outliers were discounted when using the regression tool and secondly, the parameters of the threshold model were averaged over a number of datasets.

Figure 11 shows all the models for each transform compared on a single graph. This illustrates that the diagonal subbands of the DWT and DT-CWT have significantly higher thresholds

compared to the vertical and horizontal subbands. However, the difference between the vertical and horizontal results are less pronounced within the DT-CWT. It also can be noted that the curvelet results show a pronounced reduction in thresholds compared to the other transforms (presumably from the significantly different shapes of the basis functions).

B. Discussion

For all transforms considered, the model parabola is monotonically increasing (i.e. the function exists only as the right hand side of the parabola) as shown in figures 6, 7, 9 and 10. This contrasts with the conventionally assumed minimum between 2-3 cycles/pixel as shown in figure 8. This was assumed to be because:

- there is a decrease in contrast sensitivity with increasing spatial frequency [47],
- the basis function size increases with decreasing spatial frequency.
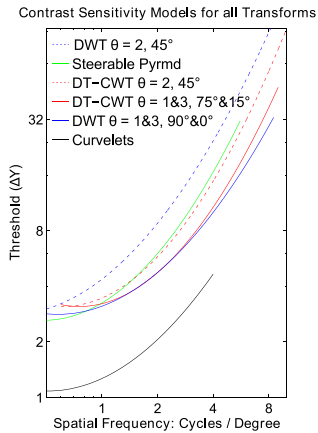
Fig. 11. Comparison of all model results (all the dotted lines from figures 6,7,9,10.

The second item has also been noted in [48] for Gabor functions similarly varying in spatial support compared to their central frequency.

The threshold function minima can also be observed to have shifted to a lower spatial frequency by at least 1.5 cycles/degree for basis function transforms that decrease in support linearly with frequency.[7] However, since these low frequencies are not likely to be analysed (i.e. over 5 levels of decomposition) in common circumstances, the considered transforms can be assumed to have monotonically increasing models.[8] This is important in many applications where the relative importance of subbands at different scales is analysed/ exploited.

1) Model Parameters: The "Oblique Effect" can be observed comparing figures 6(a), 6(b), 7(a) and 7(c) to figures 6(b) and 7(b) (although it is only significant at high frequencies as suggested in the literature). Comparing the output parameter g_2 for both the DWT and DT-CWT it can also be seen that the oblique effect (or at least the lowering of the diagonal sensitivity for high frequencies) is smaller for the DT-CWT.

Comparing our results in table 1 to those of Watson et al. [3] (also included in the table) shows that the results are broadly similar (i.e. threshold frequency minimum f_0 and "Oblique Effect" characterised by g_2 are close in value) except a different offset given by parameter a. It should be noted that the study by Watson et al. employed a significantly smaller dataset (just two observers), was conducted under different viewing conditions and used different wavelet filters. Additionally, they did not linearise their display for the majority of their experiments or use a pink noise masker between tests. Together with the fact that they also pooled spatial basis functions these differences account for the reduced thresholds found in the results of our work relative to theirs.

---

[7]This shift has been explained as the effect of higher frequencies in such basis functions contaminating lower frequency bands [49].

[8]However, tone mapping applications may require analysis of much lower frequencies than considered here and the monotonically increasing assumption may not hold in that case.

TABLE I
PARAMETERS FOR THRESHOLD MODEL FOR DWT, DT-CWT, CURVELETS AND STEERABLE PYRAMID TRANSFORMS

Transform	a	k	f_0	g_1	g_2	g_3
DWT (from [3])	0.495	0.466	0.401	1	0.534	1
DWT	2.818	0.783	0.578	1	0.555	1
DT-CWT	3.107	1.025	0.755	1	0.814	1
Curvelets	1.083	0.790	0.509	1	1	1
Steerable Pyramid	2.617	0.960	0.487	1	1	1

V. COMPARISON WITH EXISTING VISION MODELS

Many vision models exist that are able to predict detection thresholds from visual stimuli [38], [50]. Firstly, the detection thresholds of the Modelfest dataset [40] have been very well predicted using a vision model developed by Watson and Ahumada, Jr., [50]. The Modelfest dataset contains a selection of stimuli including Gabor and similar patterns (some of the stimuli resemble the elongated gratings of the Curvelet basis functions). The vision model of Watson and Ahumada, Jr., [50] has been able to accurately predict the detection thresholds of all the Modelfest stimuli. This model incorporates the effects of aperture, CSF, pooling etc. Additionally, a similar High Dynamic Range (HDR) based vision model has been developed by Mantiuk et al. [38] called HDR-VDP2. As the HDR-VDP2 system was also able to accurately model conventional Visual Dynamic Range (VDR) stimuli and because the code was freely available (as opposed to Modelfest based model [50]) we have used it to compare with our results. The basis function stimuli used in our tests were input into the HDR-VDP2 system together with a mean background "blank image" and the amplitude of the basis function stimuli was adjusted until the probability of detection P_{det} was estimated to be 0.5. The amplitude of the basis function that gave this probability of detection was considered to be the detection threshold of the stimulus (as described in section III-D and shown as ΔY) in figures 6, 7, 8, 9 and 10. The HDR-VDP2 results show a good correlation between our results and the model.[9] Specifically, these results reflect the lower frequency CSF peak of the considered basis functions in comparison to equivalent Gabor gratings.

A. HDR-VDP2 Model Error

If the mean observer thresholds t_j and the HDR-VDP2 model predictions m_j for each stimulus j are represented in dBs (i.e. in the log domain), we can use the model error RMS_{me} as defined by Watson and Ahumada, Jr., [50] to represent the HDR-VDP2 model error over all stimuli.

$$RMS_{me} = \sqrt{\frac{1}{J} \sum (t_j - m_j)^2} \qquad (6)$$

where J is the number of stimuli. For our dataset and the results shown in figures 6, 7, 9 and 10, $RMS_{me} = 3.081$ dB. This is slightly higher than the RMS_{me} value of 2.8 dB obtained by the HDR-VDP2 model using the Modelfest dataset. However, their experiments have considerable differences from our own and the standard deviation of all

---

[9]The overall sensitivity parameter of the HDR-VDP2 system was adjusted to most closely match (in a RMS sense) our data.

TABLE II

DENOISING RESULTS USING SSIM [58], VSNR [59] AND HDR-VDP2 [38] METRICS. (HDR-VDP2 PROBABILITY OF DETECTION OF DIFFERENCE BETWEEN ORIGINAL AND NOISY IMAGE; THEREFORE THE LOWER THE PROBABILITY SHOWN HERE THE BETTER THE DENOISING METHOD). DTCWT1: [57]. DTCWT2: [57] & CSF METHOD BY LASKAR ET AL. [56]. DTCWT3: [57] & OUR CSF

		Boat: 512×512 (σ)								
Metric	**Method**	**2**	**3**	**4**	**5**	**6**	**7**	**10**	**15**	**20**
SSIM	**DTCWT1**	0.9435	0.9407	0.9369	0.9332	0.9305	0.9261	0.9144	0.8956	0.8763
	DTCWT2	0.9463	0.9434	0.9397	0.9362	0.9336	0.9295	0.9120	0.9004	0.8825
	DTCWT3	**0.9529**	**0.9504**	**0.9475**	**0.9447**	**0.9430**	**0.9398**	**0.9307**	**0.9172**	**0.9012**
VSNR	**DTCWT1**	40.439	38.959	36.594	34.115	32.388	30.705	27.157	23.766	21.494
	DTCWT2	40.701	39.082	36.640	34.191	32.354	30.693	27.198	23.747	21.543
	DTCWT3	**41.373**	**39.696**	**37.018**	**34.546**	**32.582**	**30.796**	**27.267**	**23.777**	**21.606**
HDR-VDP2	**DTCWT1**	0.9554	0.9852	0.9981	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000
	DTCWT2	0.9381	0.9798	0.9974	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000
	DTCWT3	**0.8977**	**0.9663**	**0.9961**	**0.9998**	1.0000	1.0000	1.0000	1.0000	1.0000

		Lena: 512×512 (σ)								
Metric	**Method**	**2**	**3**	**4**	**5**	**6**	**7**	**10**	**15**	**20**
SSIM	**DTCWT1**	0.9652	0.9615	0.9579	0.9546	0.9520	0.9488	0.9413	0.9309	0.9201
	DTCWT2	0.9669	0.9630	0.9597	0.9565	0.9539	0.9508	0.9435	0.9333	0.9227
	DTCWT3	**0.9710**	**0.9679**	**0.9649**	**0.9621**	**0.9598**	**0.9572**	**0.9505**	**0.9400**	**0.9271**
VSNR	**DTCWT1**	40.961	39.378	37.54	35.749	34.274	33.073	30.248	27.039	24.824
	DTCWT2	41.135	39.501	37.556	35.794	34.333	33.096	30.255	**27.156**	**24.986**
	DTCWT3	**41.678**	**39.948**	**37.884**	**35.951**	**34.384**	**33.183**	**30.303**	26.996	24.876
HDR-VDP2	**DTCWT1**	0.8882	0.9702	0.9948	0.9992	1.0000	1.0000	1.0000	1.0000	1.0000
	DTCWT2	0.8410	0.9561	0.9924	0.9989	1.0000	1.0000	1.0000	1.0000	1.0000
	DTCWT3	**0.7415**	**0.9235**	**0.9878**	**0.9987**	1.0000	1.0000	1.0000	1.0000	1.0000

		Barbara: 512×512 (σ)								
Metric	**Method**	**2**	**3**	**4**	**5**	**6**	**7**	**10**	**15**	**20**
SSIM	**DTCWT1**	0.9752	0.9714	0.9673	0.9629	0.9582	0.9534	0.9396	0.9161	0.8929
	DTCWT2	0.9767	0.9733	0.9695	0.9655	0.9614	0.9570	0.9448	0.9240	0.9027
	DTCWT3	**0.9799**	**0.9771**	**0.9741**	**0.9710**	**0.9677**	**0.9644**	**0.9551**	**0.9387**	**0.9210**
VSNR	**DTCWT1**	42.228	39.343	36.963	35.014	33.477	32.075	29.123	26.011	23.695
	DTCWT2	42.422	39.488	37.112	35.170	33.650	32.138	29.339	26.147	23.957
	DTCWT3	**43.123**	**40.026**	**37.483**	**35.401**	**33.885**	**32.393**	**29.490**	**26.260**	**23.991**
HDR-VDP2	**DTCWT1**	0.9252	0.9833	0.9982	0.9998	1.0000	1.0000	1.0000	1.0000	1.0000
	DTCWT2	0.8989	0.9748	0.9970	0.9997	1.0000	1.0000	1.0000	1.0000	1.0000
	DTCWT3	**0.8443**	**0.9586**	**0.9952**	**0.9996**	1.0000	1.0000	1.0000	1.0000	1.0000

observations in our dataset is large at 4.092 dB.[10] These values together with visualisations in the above figures show that our model and results are a good fit to the visual model implemented within HDR-VDP2.

B. Relationship to Contrast Sensitivity

All results have been reported as the threshold of perception of the basis functions. In order to convert these values to contrast sensitivity S we use $S = \frac{1}{C}$ [51] where C is defined in (4).

Figure 12 compares the contrast sensitivities (S) at orientation $\theta = 1$ (75°) for both the DT-CWT basis functions and Gabor functions. This graph shows that the shape and sensitivity peak of the two functions differ significantly.[11]

[10]We have used the standard deviation of the dataset defined within [50].

[11]It should be noted that figure 12 compares the conventional Gabor-based CSF with our proposed approach: because of this, the stimuli do not have the same spatial support.

VI. APPLICATIONS

Numerous examples exist where standard CSF curves are used, together with different wavelet transforms, in image processing algorithms [52]–[56]. However, the standard CSFs used are based on Gabor grating experiments. Our work has shown that these curves are inappropriate for use with transforms such as the DWT, DT-CWT, Curvelets and the Steerable Pyramid. Therefore, when considering applications such as watermarking, compression and denoising, our precise transform based CSF / threshold information will enable exact frequency based perceptual processing.

A. Application Example: Perceptual Denoising Using the DT-CWT

Laskar et al. [56] developed a perceptual denoising algorithm that uses a standard CSF curve to modify the perceptual importance of each subband before denoising within a soft thresholding technique based
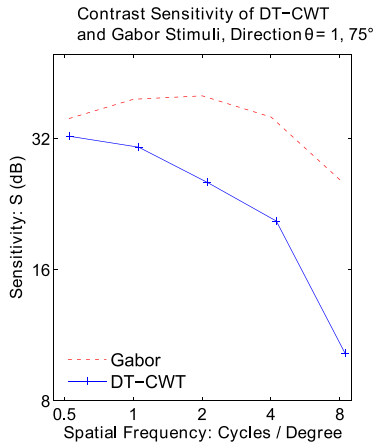
Fig. 12. CSF of Gabor filters compared to the DT-CWT.

Fig. 13. Perceptual Denoising Structure.

on a DWT decomposition. Achim and Kuruoğlu [57] have implemented a Bayesian image denoising method that utilises a bivariate Cauchy prior that exploits the cross scale statistical dependencies of the DT-CWT. As a preliminary example of how our approach can improve performance we have combined these two denoising methods to perform CSF based perceptual image denoising using the DT-CWT and the CSF/threshold models developed within our paper (figure 13). This method is identical to that used by Laskar et al. [56] but replaces the DWT with the DT-CWT and uses a bivariate thresholding technique [57] (the DT-CWT has considerable advantages over the DWT as discussed in [57]).

1) Results: Table II shows the results of denoising experiments where variable amounts of white noise (with standard deviation σ) are added to three test images and the results are assessed using SSIM [58], VSNR [59] and HDR-VDP2 [38]. The SSIM metric is a standard perceptually based metric that measures the structural similarity of two images [58]. However, SSIM does not explicitly use any definition of a CSF and therefore we have also included the VSNR and HDR-VDP2 metrics which do. The HDR-VDP2 output gives the probability of detection of a difference between the original and the noisy / denoised image. Therefore (unlike SSIM and VSNR) smaller output values indicate better performing methods. Results for low amounts of noise ($2 \leq \sigma \leq 7$) are included as the probabilities of detection generated by VDP-HDR2 are only meaningful within this range (for noise levels where $\sigma \geq 10$ the distortion is obvious reflecting a probability of 1.0 for all methods).

This table shows:

- **DTCWT1:** The results of the bivariate denoising method using the DTCWT transform developed by Achim and Kuruoğlu [57].
- **DTCWT2:** Identical to DTCWT1 but uses the perceptual weighting method shown in figure 13 in conjunction with the CSF curve used by Laskar et al. [56].
- **DTCWT3:** Identical to DTCWT2 but uses the CSF/threshold results presented in our paper.

The images were all of resolution 512×512 pixels and greyscale. The DT-CWT used the near symmetric first level filters [39] and the transform contained 6 levels.
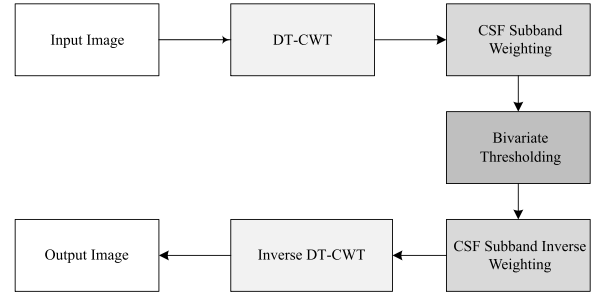
The results indicate that the integration of CSF curves used by Laskar et al. [56] (DTCWT2) provides better results compared to the original method by Achim et al. (DTCWT1). However, by using our new CSF/threshold information (DTWCWT3) considerable improvement in the results are shown when compared to both DTCWT1 and DTCWT2. The improved results of DTCWT3 over DTCWT2 demonstrates that conventional CSF curves are not appropriate for image processing with analysis transforms such as the DTCWT.

VII. Conclusion

We have presented the perceptual thresholds and contrast sensitivities of a range of basis functions from popular wavelet analysis decompositions (the DWT, Steerable Pyramid, DT-CWT and Curvelet transforms). Whereas typical CSF graphs show a peak of sensitivity at approximately 2 cycles/degree, we have demonstrated that only high frequency attenuation of sensitivity is apparent for the considered basis functions over common frequency ranges. The lack of low frequency attenuation is assumed to be caused by the variation in basis function support (low frequency basis functions having larger support than high frequency basis functions). Although a previous study has shown this effect for the DWT, our study is the first to be able to confirm similar trends in the results for the range of more flexible transforms. The results were shown to be comparable to the thresholds generated from the same basis functions using the visual model HDR-VDP2. Similar variations in CSF shape have been found through the variation of grating area by Rovamo *et al.* [60]. However, the basis functions we have considered are not simple Gabor or gratings and therefore our results are precisely realised for the considered transforms.

These new models have demonstrated quantitative improvements in an example denoising application. Our quantitative results facilitate the next generation of perceptual processing applications when using any of the analysed transforms. MATLAB code giving precise thresholds and contrast sensitivities for all subbands for all the transforms are publicly available on figshare https://shar.es/1urZ5a.

REFERENCES

[1] Z. Liu, L. J. Karam, and A. B. Watson, "JPEG2000 encoding with perceptual distortion control," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1763–1778, Jul. 2006.

[2] Y. Zhang, E. Reinhard, and D. Bull, "Perception-based high dynamic range video compression with optimal bit-depth transformation," in *Proc. 18th IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 1321–1324.

[3] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," *IEEE Trans. Image Process.*, vol. 6, no. 8, pp. 1164–1175, Aug. 1997.

[4] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 37, no. 12, pp. 2091–2110, Dec. 1989.

[5] N. Kingsbury, "A dual-tree complex wavelet transform with improved orthogonality and symmetry properties," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2000, pp. 375–378.

[6] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 123–151, Nov. 2005.

[7] A. B. Watson, "The cortex transform: Rapid computation of simulated neural images," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 311–327, Sep. 1987.

[8] A. Lukin, "Improved visible differences predictor using a complex cortex transform," in *Proc. Int. Conf. Comput. Graph. Vis. (GraphiCon)*, 2009, pp. 1–6.

[9] E. H. Adelson, E. Simoncelli, and R. Hingorani, "Orthogonal pyramid transforms for image coding," in *Proc. Cambridge Symp. Int. Soc. Opt. Photon.*, 1987, pp. 50–58.

[10] M. Miller and N. Kingsbury, "Image modeling using interscale phase properties of complex wavelet coefficients," *IEEE Trans. Image Process.*, vol. 17, no. 9, pp. 1491–1499, Sep. 2008.

[11] T. Wan, N. Canagarajah, and A. Achim, "Segmentation of noisy colour images using cauchy distribution in the complex wavelet domain," *IET Image Process.*, vol. 5, no. 2, pp. 159–170, Mar. 2011.

[12] P. Hill, N. Canagarajah, and D. Bull, "Image fusion using complex wavelets," in *Proc. 13th Brit. Mach. Vis. Conf.*, 2002, pp. 487–496.

[13] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861–899, 2006.

[14] A. S. Gaikwad and D. S. Bormane, "Image denoising using digital image curvelet," *J. Inf. Eng. Appl.*, vol. 1, no. 2, pp. 23–28, 2011.

[15] J.-L. Starck, E. J. Candès, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Trans. Electron. Packag. Manuf.*, vol. 11, no. 6, pp. 670–684, Jun. 2002.

[16] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 9, pp. 891–906, Sep. 1991.

[17] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in *Proc. Int. Conf. Image Process.*, vol. 3. Oct. 1995, pp. 444–447.

[18] S. Daly, "The visible differences predictor: An algorithm for the assessment of image fidelity," in *Digital Images and Human Vision*. Cambridge, MA, USA: MIT Press, 1993, pp. 179–206.

[19] P. G. J. Barten, *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*, vol. 72. Bellingham, WA, USA: SPIE, 1999, pp. 2759–2770.

[20] A. van Meeteren and J. J. Vos, "Resolution and contrast sensitivity at low luminances," *Vis. Res.*, vol. 12, no. 5, pp. 825–833, 1972.

[21] K. J. Kim, R. Mantiuk, and K. H. Lee, "Measurements of achromatic and chromatic contrast sensitivity functions for an extended range of adaptation luminance," *Proc. SPIE Human Vis. Electron. Imag. XVIII*, vol. 8651, p. 86511A, Mar. 2013.

[22] F. L. Van Nes, J. J. Koenderink, H. Nas, and M. A. Bouman, "Spatiotemporal modulation transfer in the human eye," *J. Opt. Soc. Amer.*, vol. 57, no. 9, pp. 1082–1088, 1967.

[23] P. J. Bex, S. G. Solomon, and S. C. Dakin, "Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure," *J. Vis.*, vol. 9, no. 10, pp. 401–406, 2009.

[24] D. M. Chandler and S. S. Hemami, "Dynamic contrast-based quantization for lossy wavelet image compression," *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 397–410, Apr. 2005.

[25] G. M. Johnson and M. D. Fairchild, "On contrast sensitivity in an image difference model," in *Proc. IS&T PICS Conf.*, 2001, pp. 18–23.

[26] A. B. Watson and J. A. Solomon, "Model of visual contrast gain control and pattern masking," *J. Opt. Soc. Amer. A*, vol. 14, no. 9, pp. 2379–2391, Sep. 1997.

[27] M. A. Georgeson and G. D. Sullivan, "Contrast constancy: Deblurring in human vision by spatial frequency channels," *J. Physiol.*, vol. 252, no. 3, pp. 627–656, 1975.

[28] J. Fiser, P. J. Bex, and W. Makous, "Contrast conservation in human vision," *Vis. Res.*, vol. 43, no. 25, pp. 2637–2648, 2003.

[29] S. J. Daly, "Application of a noise-adaptive Contrast Sensitivity Function to image data compression," *Opt. Eng.*, vol. 29, no. 8, pp. 977–987, 1990.

[30] B. C. Hansen, A. M. Haun, and E. A. Essock, "The 'horizontal effect': A perceptual anisotropy in visual processing of naturalistic broad-band stimuli," in *Visual Cortex: New Research*. New York, NY, USA: Nova Science Publishers, 2008.

[31] A. van der Schaaf and J. H. van Hateren, "Modelling the power spectra of natural images: Statistics and information," *Vis. Res.*, no. 4, pp. 2759–2770, Sep. 1996.

[32] B. C. Hansen and E. A. Essock, "A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes," *J. Vis.*, vol. 4, pp. 1044–1060, Dec. 2004.

[33] F. W. Campbell, J. J. Kulikowski, and J. Levinson, "The effect of orientation on the visual resolution of gratings," *J. Physiol.*, vol. 187, no. 2, pp. 427–436, 1966.

[34] C. M. Murasugi and P. Cavanagh, "Anisotropy in the chromatic channel: A horizontal-vertical effect," *Spatial Vis.*, vol. 3, no. 4, pp. 281–291, 1988.

[35] D. H. Brainard, "The psychophysics toolbox," *Spatial Vis.*, vol. 10, no. 4, pp. 433–436, 1997.

[36] I. Höntsch and L. J. Karam, "Adaptive image coding with perceptual distortion control," *IEEE Trans. Image Process.*, vol. 11, no. 3, pp. 213–222, Mar. 2002.

[37] H. A. Peterson, A. J. Ahumada, Jr., and A. B. Watson, "Improved detection model for DCT coefficient quantization," in *Proc. SPIE IS&T/SPIE's Symp. Electron. Imag., Sci. Technol.*, Sep. 1993, pp. 191–201.

[38] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.*, vol. 30, no. 4, 2011, Art. no. 40.

[39] A. F. Abdelnour and I. W. Selesnick, "Nearly symmetric orthogonal wavelet bases," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2001, pp. 344–349.

[40] T. Carney *et al.*, "Development of an image/threshold database for designing and testing human vision models," in *Proc. SPIE Electron. Imag.*, May 1999, pp. 542–551.

[41] Y. Zhang, D. Agrafiotis, M. Naccari, M. Mrak, and D. R. Bull, "Visual masking phenomena with high dynamic range content," in *Proc. IEEE ICIP*, Sep. 2013, pp. 2284–2288.

[42] A. B. Watson and D. G. Pelli, "QUEST: A Bayesian adaptive psychometric method," *Perception Psychophys.*, vol. 33, no. 2, pp. 113–120, 1983.

[43] A. B. Watson, "Probability summation over time," *Vis. Res.*, vol. 19, no. 5, pp. 515–522, 1979.

[44] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *J. Opt. Soc. Amer. A*, vol. 4, no. 12, pp. 2379–2394, Dec. 1987.

[45] T. Kunkel and E. Reinhard, "A reassessment of the simultaneous dynamic range of the human visual system," in *Proc. ACM 7th Symp. Appl. Perception Graph. Visualizat.*, Jul. 2010, pp. 17–24.

[46] F. E. Grubbs, "Procedures for detecting outlying observations in samples," *Technometrics*, vol. 11, no. 1, pp. 1–21, 1969.

[47] F. L. van Nes and M. A. Bouman, "Spatial modulation transfer in the human eye," *J. Opt. Soc. Amer.*, vol. 57, no. 3, pp. 401–406, 1967.

[48] A. B. Watson, "Estimation of local spatial scale," *J. Opt. Soc. Amer. A*, vol. 4, no. 8, pp. 1579–1582, 1987.

[49] E. Peli, L. E. Arend, G. M. Young, and R. B. Goldstein, "Contrast sensitivity to patch stimuli: Effects of spatial bandwidth and temporal presentation," *Spatial Vis.*, vol. 7, no. 1, pp. 1–14, 1993.

[50] A. B. Watson and A. J. Ahumada, Jr., "A standard model for foveal detection of spatial contrast," *J. Vis.*, vol. 5, no. 9, p. 6, 2005.

[51] F. W. Campbell and D. G. Green, "Optical and retinal factors affecting visual resolution," *J. Physiol.*, vol. 181, no. 3, pp. 576–593, 1965.

[52] M. J. Nadenau, J. Reichel, and M. Kunt, "Wavelet-based color image compression: Exploiting the contrast sensitivity function," *IEEE Trans. Image Process.*, vol. 12, no. 1, pp. 58–70, Jan. 2003.

[53] H. Song and J. Gu, "Curvelet based adaptive watermarking for images," in *Proc. 2nd Int. Conf. Comput. Sci. Netw. Technol. (ICCSNT)*, Dec. 2012, pp. 1101–1105.

[54] J. N. Ellinas and P. Kenterlis, "A wavelet-based watermarking method-exploiting the contrast sensitivity function," *Int. J. Signal Process.*, vol. 3, no. 4, pp. 266–272, 2006.

[55] K.-Q. Huang, Z.-Y. Wu, G. S. K. Fung, and F. H. Y. Chan, "Color image denoising with wavelet thresholding based on human visual system model," *Signal Process., Image Commun.*, vol. 20, no. 2, pp. 115–127, 2005.

[56] H. H. Laskar, S. Baishya, S. K. Kar, R. Sharma, N. Medhi, and R. D. Purkayastha, "Color image denoising in wavelet domain using adaptive thresholding incorporating the human visual system model," in *Proc. IEEE Int. Conf. Elect. Comput. Eng. (ICECE)*, Dec. 2010, pp. 498–501.

[57] A. Achim and E. E. Kuruoğlu, "Image denoising using bivariate α-stable distributions in the complex wavelet domain," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 17–20, Jan. 2005.

[58] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[59] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.

[60] J. Rovamo, O. Luntinen, and R. Näsänen, "Modelling the dependence of contrast sensitivity on grating area and spatial frequency," *Vis. Res.*, vol. 33, no. 18, pp. 2773–2788, 1993.

Alin Achim received the M.Eng. and M.Sc. degrees in electrical engineering from the Politehnica University of Bucharest, Romania, in 1995 and 1996, respectively, and the Ph.D. degree in biomedical engineering from the University of Patras, Greece, in 2003. He received the European Research Consortium for Informatics and Mathematics Postdoctoral Fellowship which he spent with the Institute of Information Science and Technologies, Pisa, Italy, and with the French National Institute for Research in Computer Science and Control, Sophia Antipolis, France. He is currently a Reader in biomedical image computing with the Department of Electrical and Electronics Engineering, University of Bristol, U.K. He joined Bristol in 2004, as a Lecturer, and then became a Senior Lecturer (Associate Professor) in 2010. He has co-authored over 100 scientific publications, including 30 journal papers. His research interests are in statistical signal, image and video processing, multiresolution algorithms, image filtering, fusion, and super-resolution, segmentation and classification. He is member of the IET and a Senior Member of the IEEE. He is an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, and an Elected Member of the Bio Imaging and Signal Processing Technical Committee of the IEEE Signal Processing Society and an Affiliated Member of the Society's Signal Processing Theory and Methods Technical Committee.

Mohammed Ebrahim Al-Mualla received the B.Eng. degree in telecommunications from the Etisalat College of Engineering, UAE, and the M.Sc. degree in communication systems and signal processing and the Ph.D. degree in electrical and electronics engineering from The University of Bristol, U.K. He is currently the Senior Vice President of Research and Interim Provost with Kahlifa University, UAE.

David Bull (M'94–SM'07–F'12) received the B.Sc. degree from the University of Exeter, in 1980, the M.Sc. degree from the University of Manchester, in 1983, and the Ph.D. degree from the University of Cardiff, in 1988. He currently holds the Chair in Signal Processing with The University of Bristol, where he was the Head of the Electrical and Electronic Engineering Department from 2001 to 2006. He heads the Visual Information Laboratory and is the Director of the Bristol Vision Institute. In 2001, he co-founded ProVision Communication Technologies, Ltd. His current activities are focused on the problems of image and video communications and analysis for wireless, Internet, and broadcast applications. He has authored over 400 academic papers.

Paul Hill received the B.Sc. degree from Open University, in 1996, and the M.Sc. degree and the Ph.D. degree in image segmentation and fusion from The University of Bristol, Bristol, U.K., in 1998 and 2002, respectively. His research interests include image and video analysis, compression, and fusion. He is currently a Senior Research Fellow with the Department of Electrical and Electronic Engineering, The University of Bristol, also lecturing in audio technology. He has authored over 25 academic papers.