

Visually Significant Edges

Tunç Ozan Aydın*
MPI Informatik

Martin Čadík†
MPI Informatik

Karol Myszkowski‡
MPI Informatik

Hans-Peter Seidel§
MPI Informatik

Abstract

Numerous computer graphics methods make use of either explicitly computed strength of image edges, or an implicit edge strength definition that is integrated into their algorithms. In both cases, the end result is highly affected by the computation of edge strength. We address several shortcomings of the widely used gradient magnitude based edge strength model through the computation of a hypothetical human visual system (HVS) response at edge locations. Contrary to gradient magnitude, the resulting “visual significance” values account for various HVS mechanisms such as luminance adaptation and visual masking, and are scaled in perceptually linear units that are uniform across images. The visual significance computation is implemented in a fast multi-scale second generation wavelet framework, which we use to demonstrate the differences in image retargeting, HDR image stitching and tone mapping applications with respect to gradient magnitude model. Our results suggest that simple perceptual models provide qualitative improvements on applications utilizing edge strength at the cost of a modest computational burden.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation; I.4.6 [Segmentation]: Edge and Feature Detection; I.4.2 [Enhancement]: Filtering

Keywords: edge strength, visual perception, HDR

1 Introduction

Localizing significant variations in image luminance and chrominance, i.e. edge detection, has been a classical problem in image processing. Similarly, edge aware image decompositions have been used in numerous computer graphics applications such as image abstraction, detail enhancement and HDR tone mapping. In both contexts, the essential component is an edge model, which in the former case is used to produce a map of image edges, and in the latter case is integrated into the image decomposition algorithm that purposely avoids smoothing near strong edges.

The edge model serves two purposes: determining the location and strength of edges. The majority of the methods proposed for edge detection involve smoothing and differentiation to locate edges. A measure of edge strength is essential, since typically the result of these methods is “too many” edges, and the output is only comprehensible after the removal “less important” edges thorough thresholding. Incidentally, gradient magnitude based edge models are conveniently used in all but the most specialized edge detectors, because one can locate edges by computing local maxima of the gradient magnitude, as well as simply use the magnitude value at the edge location as a rough estimate of edge strength.

While existing methods are capable of localizing edges in a semantically meaningful way, their performance is directly influenced by the edge strength model they employ. The focus of this work is the

computation of edge strength rather than edge localization and semantics. Our central idea is that the magnitude of image edges as perceived by the human eye, or the “visual significance” of an edge, should be the guideline for edge strength computation. In that respect, gradient magnitude as an edge strength measure encapsulates the well known property of the Human Visual System (HVS) being sensitive to luminance differences, but ignores other aspects such as visual masking and luminance adaptation. Earlier research [Ferwerda et al. 1997] has demonstrated how image contrast is masked by other contrast patches that are of similar spatial frequencies. Except perhaps simple stimuli designed for experimental purposes, visual masking is expected to occur in virtually any complex image and often to have a strong influence on perception. Disregarding the non-linear perception of luminance, especially in HDR images, often leads to overestimations in bright image regions. As a simple counter-measure, one can operate in log-luminance space [Fattal et al. 2002] that better approximates perceived intensity in bright image regions, but fails to model the perception of lower luminance values that is not linear in log-space.

We present an edge aware image decomposition framework based on second generation wavelets [Fattal 2009] that uses visual significance as its edge strength metric. *The contribution of this work is the use of an HVS model to estimate visual significance as a measure of edge strength*, instead of gradient magnitude that is commonly used in computer graphics applications. The HVS model computes physical contrast at edge locations, and scales it through a cascade of simple and well known models of luminance adaptation, spatial frequency perception and visual masking. The computed visual significance is approximately scaled in perceptually linear units, which implies that similar edge strength values across multiple images correspond to similar perceived strengths. In this paper, we first summarize related work (Section 2), then discuss the edge aware decomposition framework (Sec. 3) and the HVS model (Section 4), then we validate the model (Sec. 5) and show that the use of visually significant edges results in qualitatively better outcomes in image retargeting, panorama stitching and HDR tone mapping over gradient magnitude based approaches (Sec. 6).

2 Background

In this section we discuss related work on edge detection, computer graphics applications that utilize edge models, and HVS models for contrast perception. Due to the purely 2D nature of our technique, we do not discuss any line drawing techniques that are capable of localizing edges in a semantically meaningful way, but require 3D information about depicted objects.

Edge Detection

Edge Detection has been one of the fundamental problems in computer vision. In an early approach, Marr and Hildreth used the zero crossings of the Laplacian operator motivated by its rotational symmetry [Marr and Hildreth 1980]. Later Canny focused on finding an optimal differential operator that localizes sharp intensity edges (which he approximated with the first derivative of a Gaussian), and introduced the use of non-maxima suppression and hysteresis thresholding [Canny 1986]. Canny’s method proved to be very reliable over the years and is still widely used. A notable improvement over earlier edge detectors is the use of multi-scale analysis to detect smooth edges as well as sharper edges (see [Pellegrino

*e-mail: tunc@mpi-inf.mpg.de

†mcardik@mpi-inf.mpg.de

‡karol@mpi-inf.mpg.de

§hpseidel@mpi-inf.mpg.de

et al. 2004] for an overview). The steerable pyramid decomposition, while designed for general purpose feature detection, is shown to perform better at small peaks of intensity by combining even and odd filter responses [Freeman and Adelson 1991]. Lindeberg proposed an automatic scale selection method where the scale of edges is determined by finding the maximum of a strength measure over scales [Lindeberg 1996]. This method is later employed in Georgeson’s third derivative operator [Georgeson et al. 2007], which provides a more compact response than the first derivative. Some effort has also been made to detect color edges [Ruzon and Tomasi 1999]. For a detailed summary of edge detection techniques we refer the reader to [Ziou and Tabbone 1997].

Applications

Edge detection has found various applications in computer graphics such as guidance over image editing operations [Elder and Goldberg 2001], stylization and abstraction of photographs [DeCarlo and Santella 2002] and texture flattening [Perez et al. 2003]. The notion of edge importance understood as its “lifetime” (essentially its presence) over increasing scales in the scale-space framework similar to [Lindeberg 1996] has been used for stylized line drawings and structure-aware image abstraction [Orzan et al. 2007]. Edge-preserving techniques such as the bilateral filter have been used to decompose an image into a base and detail layers and applied to HDR tone mapping [Durand and Dorsey 2002]. Recently, Farbman et al. [Farbman et al. 2008] proposed another decomposition with multiple detail layers and presented applications to scale selective feature enhancement and image abstraction. Fattal [2009] later showed that comparable results can be achieved much faster using a second generation wavelet decomposition with a specialized weighting function that avoids edges. Another approach to edge preserving filtering is detecting the edge strength by computing the gradient of the input image, and reconstructing the image through anisotropic diffusion [Perona and Malik 1990]. This method decouples edge detection and smoothing, but it is inefficient due to the iterative processing. This method has later been modified by an edge strength measure based on curvature change [Tumblin and Turk 1999]. Gradient domain operators such as [Fattal et al. 2002; Mantiuk et al. 2006], while not explicitly stated, also utilize edges since gradient magnitude operator is essentially an edge detector. Mantiuk et al.’s [2006] method has additionally a perceptual component in the form of a simple contrast transducer.

Contrast Perception

The HVS characteristics involved in contrast perception are quite complex and have been investigated in numerous psychophysical studies. Even in the simple case of *detection experiments*, where the task is to distinguish a sine wave grating from the uniform background, the resulting detection threshold depends on many factors such as the background (adaptation) luminance, the grating’s spatial frequency, orientation, spatial extent, and eccentricity with respect to the fovea. These characteristics are modeled by contrast sensitivity functions (CSF) [Daly 1993; Barten 1999]. Other characteristics of contrast perception are observed in the *discrimination experiments*, whose goal is to determine how the presence of one masking sine [Legge and Foley 1980] or square [Whittle 1986] grating affects the discriminability of another test grating. In some experiments, it turned out that the maskers of weak contrast actually facilitate the discriminability of test grating, and the corresponding discrimination thresholds are even smaller than the detection threshold as measured by the CSF. For high contrast (suprathreshold) maskers an elevation of discrimination thresholds can be observed. This behavior is modeled by *transducer* functions [Legge and Foley 1980; Wilson 1980; Mantiuk et al. 2006], which convert physical contrast of an image to a hypothetical HVS response. Various transducers have been successfully incorporated into the HVS models used in many computer graphics applications including texture mask-

ing simulation [Ferwerda et al. 1997], image appearance modeling [Pattanaik et al. 1998], perception-based rendering [Bolin and Meyer 1998], and tone mapping and contrast enhancement [Mantiuk et al. 2006; Mantiuk et al. 2008]. Often, transducer functions limit their modeling to intra-channel masking assuming a certain contrast patch is solely masked by other contrast patches at the same spatial frequency and orientations. A more comprehensive model by Watson and Solomon [Watson and Solomon 1997] also comprises masking from adjacent frequencies (inter-channel masking), in effect contrast patches are subject to masking from other contrast patches within a certain neighborhood. The neighborhood masking model in JPEG2000 is a simpler implementation of the same principle [Zeng et al. 2000].

3 Edge Avoiding Framework

Objects appear differently depending on the scale of observation, and thus visual significance of image features depends on the image scale. Consequently, many image processing tools including edge detection algorithms adopted multi-scale approaches. This has been physiologically justified by the finding that each simple retinal cell responds to a certain bandwidth of spatial frequencies [Wandell 1995, Chapter 6].

Recent work [Fattal 2009] demonstrates use of second generation wavelets computed through the lifting scheme [Sweldens 1997] in the context of edge avoiding multi-scale image decomposition. In this section we give an overview of these concepts, for a detailed discussion refer to [Jansen and Oonincx 2005]. Contrary to regular wavelets, second generation wavelet bases do not have to be merely translates and dilates of a single pair of scaling and wavelet functions. This generalization enables data dependent filtering through the use of a weighting function that utilizes the information obtained from the local neighborhood changes the shape of wavelet bases accordingly. In the context of edge avoiding wavelets (EAW) the weighting function assigns lower weights to locations containing strong edges, thus the wavelet bases effectively “avoid” those locations.

The data dependent filtering achieved by wavelet bases not relying on translation and dilation comes at the cost of prohibiting the use of Fourier analysis for wavelet calculation. This issue has been addressed by a discrete wavelet transform named the lifting scheme [Sweldens 1997]. The basic idea behind the lifting scheme is to *split* a signal into fine and coarse samples, *predict* fine samples from coarse samples and compute the details by subtracting fine samples from their prediction, and *update* coarse samples using the details. Fig. 1 illustrates the computation in 1D (using Uytterhoeven’s coloring scheme [Uytterhoeven et al. 1997]). Advantages of the lifting scheme are fast, in place computation and easily invertible decomposition.

One can achieve edge aware behavior by simply executing a weighting function at each location that assigns weights according to the edge strength at the local neighborhood. If the goal is to avoid edges, i.e. obtaining detail components free of strong edges, this can be achieved by the function ω in Equation 1, where m and n are intensities at the current location and some neighboring pixel, respectively:

$$\omega(m, n) = \frac{1}{(|\nu(m, n)|^\alpha + \epsilon)}. \quad (1)$$

The control parameter α is set to 0.8 as suggested in [Fattal 2009]. Divisions by zero are prevented by setting ϵ to 10^{-5} . We will use the function ν later for the estimation of visual significance; in the original implementation it simply returns the difference of n and m . Such a decomposition is useful in contrast editing applications

such as detail enhancement and image abstraction, since halo artifacts are prevented due to the absence of strong edges in detail components. The opposite goal of extracting solely strong edges can be achieved by simply using the *inverse* of ω . The detail components of the resulting decomposition closely resemble the outcome of multi-scale edge detectors, which we utilize in context aware image retargeting and panorama stitching applications (Section 6).

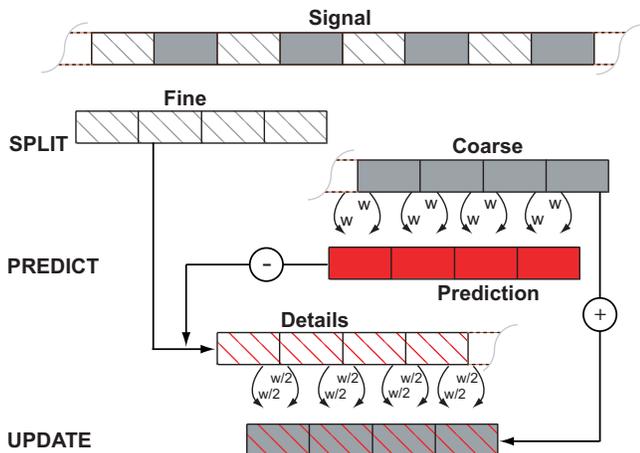


Figure 1: Illustration of the lifting scheme on a 1D signal. The signal is decomposed into fine and coarse parts by designating odd pixels as fine, and even pixels as coarse components. The fine component is predicted from the coarse component using weights computed by the edge aware function ω , or simply by linear interpolation. The difference between the original fine component and the predicted fine component gives the details. The details are then used to update the coarse component. The same process is then iterated on the updated coarse signal.

The straightforward extension to the second dimension is to repeat the 1D computation at both dimensions (Fig. 2a). If an edge preserving weighting function is used, the results of this 2D decomposition are analogous to X and Y gradients, and thus fit naturally into the edge detection pipeline. Another splitting method by [Uytterhoeven et al. 1997] with lower anisotropy produces better results coupled with an edge avoiding weighting function (Fig. 2b).

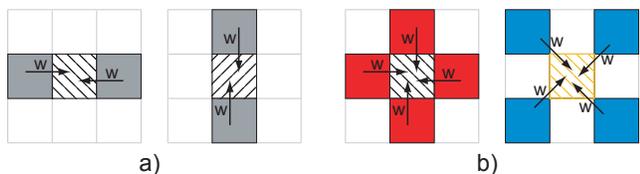


Figure 2: The lifting scheme can be extended by repeating the 1D computation in X and Y directions (a), or using a lower anisotropy red-black quincunx lattice (b). Only the prediction step is illustrated for brevity.

4 Human Visual System Model

We extend the EAW framework (Section 3) with an HVS model, where we modify the weighting function (Equation 1) that penalizes strong differences of image pixel values by computing visual

significance of the luminance differences. The HVS model takes physical image luminance as input, therefore 8-bit images should be mapped to display luminance and HDR images should be calibrated to scene luminance before processing. The luminance **contrast** C is approximated in the EAW framework by dividing the fine samples by the local mean of the *predictions* of immediate neighbors K (2 and 4 for X-Y splitting and red-black splitting, respectively):

$$C = \frac{\text{Fine}}{\left(\frac{1}{K}\right) \sum_K \text{Prediction}_k} - 1. \quad (2)$$

Repeated at each scale, this formulation is similar to the low-pass contrast in [Mantiuk et al. 2006]. The advantage of a contrast based edge strength measure over a gradient based measure is illustrated in Fig.3

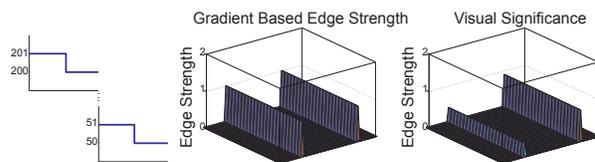


Figure 3: Edge strength predictions utilizing physical contrast account for the effect of background luminance level. The perceived strength of step edges 200-201 cd/m^2 and 50-51 cd/m^2 (left) are predicted to be the same by the gradient based method, whereas a contrast based method correctly predicts the weaker perceived strength of the first profile.

Note that the contrast C is computed solely using physical luminance. As the next step we scale C by computing the sensitivity of the visual system to obtain contrast in perceptually linear units. Two prominent factors that affect **contrast sensitivity** are its spatial frequency (ρ), and the adaptation luminance (L_a). These effects can easily be observed in the Campbell-Robson chart. We use the CSF from the Visible Differences Predictor [Daly 1993] with corrections as indicated in [Aydın et al. 2008, Equations (10, 11)] to obtain the perceptually linearized contrast $C' = C \cdot \text{CSF}(\rho, L_a)$. Fig. 4 shows an example where the difference in edge preserving smoothing is mainly due to the scaling of contrast by the CSF. This behavior is typical in HDR images, where the contrast magnitudes at very bright and very dark image regions are overestimated by the frameworks without perceptual components. As a result, the edges of the bright window are avoided unlike the edges at the window's frame (Fig. 4 center). The CSF's scaling results in a more uniform smoothing over edges with similar magnitude of visibility (Fig. 4 right).

Visual masking is the decrease in visibility of a contrast patch in the presence of other contrast patches of similar spatial frequencies. One way of modeling this effect is by computing a *threshold elevation* map for each visual channel, which when divided by the contrast at that channel accounts for the increase in detection thresholds (thus, decrease in sensitivity). This method trades off accuracy at supra-threshold contrast levels for better prediction near the threshold, and has been used in image quality assessment metrics for distortion detection. On the other hand, the *transducer* model is focused on perception of supra-threshold contrasts and thus preferred in discrimination tasks. The model relies on a transducer function that is constructed by iteratively summing up contrast detection thresholds. The use of a transducer function in computer graphics context is demonstrated in [Ferwerda et al. 1997]. A more comprehensive transducer model [Watson and Solomon 1997] also comprises masking from adjacent frequency channels (inter-channel masking). In this model, since the lower frequency



Figure 4: The effect of luminance adaptation. The original HDR image (left), smoothing with EAW method (center), and smoothing with EAW method using visually significant edges (right). The strength of edges of the bright window are overestimated by EAW method in the absence of a model of luminance adaptation. All images are tone mapped [Reinhard et al. 2002] for display purposes.

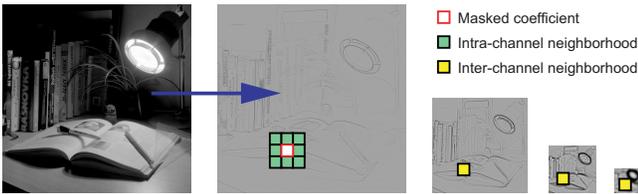


Figure 5: An illustration of neighborhood masking on detail layers of a multi-scale decomposed image.

channels contain information from the spatial neighborhood, a contrast patch at a certain location is effectively masked by neighboring contrast patches (See Fig. 5 for an illustration of neighborhood masking.)

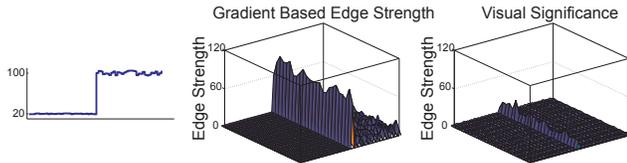


Figure 6: The visual masking due to the random noise modulated by image luminance in the test stimulus (left), results in lower perceived edge strength than the gradient magnitude (center), as predicted by our method (right).

While the visual masking due to the local neighborhood is often not significant for isolated test stimuli, natural images tend to have “busy”, textured regions where the visibility of edges are notably lesser than non-textured regions. To account for that, our ν function (Equation 1) comprises the point-wise extended masking model [Zeng et al. 2000] which, in addition to a compressive non-linearity, also accounts for visual masking from the local neighborhood K :

$$R = \frac{\text{sign}(C')|C'|^{0.5}}{(1 + \sum_K |C'_k|^{0.2})}. \quad (3)$$

The effect of visual masking on a simple stimulus is illustrated in Fig. 6. Figure 7 shows that the involvement of the point-wise extended masking model results in a perceptually uniform smoothing near high-masking regions. Computation of the hypothetical HVS response R is the final step in function ν in EAW the framework.



Figure 7: The effect of contrast masking in a complex image. The original image (left), smoothing with EAW method (center), and smoothing with EAW method using visually significant edges (right). The masking model reduces the strength of the facial hair edges due to the presence of hair in the local neighborhood.

5 Model Calibration – Perceptual Experiment

To validate and calibrate the proposed edge perception model, we conducted a simple threshold-level perceptual experiment. The motivation for this is twofold: first, we aim to calibrate the implemented supra-threshold transducer model described above (Equation 3) for threshold stimuli; second, as noted by [Whittle 1986], discrimination thresholds for spatially separated patches should not be generalized for perceiving edges, thus there is a lack of usable experimental data. Furthermore, the used CSF curves [Daly 1993] reflect measurements using the Michelson’s definition of contrast, which is slightly different from the implemented definition contrast (Equation 2).

In our experiment, two adjacent grayscale patches were presented on a calibrated display device. The luminance of the left patch is kept constant during each trial, whereas the luminance of the right patch was modulated according to the responses of the subject. Each subject was asked whether there is a visible edge between the two patches or not. The luminance of the right patch was decreased if the response was positive, and increased if the response was negative. The step sizes were determined by following the PEST procedure [Taylor and Creelman 1967]. A random noise pattern was presented for 1s between stimuli to avoid after-images, memory effects, etc. Each trial ended once the standard deviation of the subject’s last 6 responses were below the minimum step size (0.01 cd/m^2) or if there were more than 30 responses collected. The experiment comprised 10 trials for each subject, where the initial luminance of the left patch at each trial is selected by randomized sampling from the luminance range $1.5 - 400 \text{ cd/m}^2$.

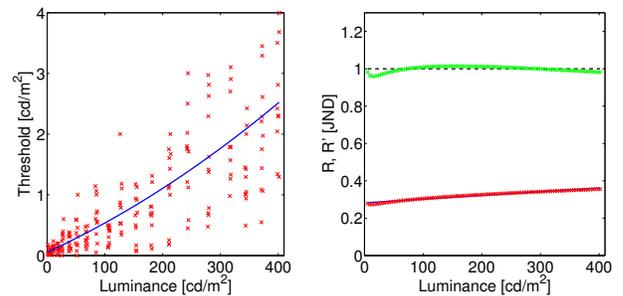


Figure 8: Perceptual experiment. Left: measured edge detection luminance thresholds as a function of adaptation luminance L_a , right: model predictions before (red crosses) and after the calibration (green crosses). An ideal model response is constantly 1 JND for the threshold data (dashed line).

The stimuli were displayed on a calibrated Barco Coronis MDCC

3120 DL, a 10-bit 21-inch hi-precision LCD display, in its native resolution 2048×1536 pixels, the maximal display luminance was 440 cd/m^2 . The display response was measured by the Minolta LS-100 luminance meter. The experimentation room was darkened (measured light level: 1 lux), and observers sat approximately 70 cm from the display. The total of 22 observers took part in our experiment. The observers were both male and female and all of them reported to have normal, or corrected-to-normal vision. Each subject was verbally introduced to the problem before the experiment.

The measured edge perception thresholds, see Fig. 8 (left), were approximated by the second order polynomial function (blue curve). Using the polynomial function, we generated 100 input threshold stimuli as the inputs for model calibration procedure. We assume that the model output for each stimulus at the threshold level should be $R=1$ JND. Therefore, we run the model for each of 100 input stimuli to obtain the error function, see Fig. 8 (right). The threshold prediction of the uncalibrated model (red crosses) was quite solid, so that we decided to perform the calibration by means of a simple linear function which should not affect the performance of the model for supra-threshold stimuli. The calibration was achieved by dividing the masking model by the calibration function (blue curve in Fig. 8 (right)):

$$R' = \frac{R}{0.0002 L_a + 0.2822}, \quad (4)$$

where L_a is the adaptation luminance in cd/m^2 .

As the masking model (Equation 3) was verified in JPEG 2000 applications, we did not calibrate it for supra-threshold data. However, we believe that the supra-threshold performance is also improved as a consequence of the threshold calibration, and the precision of the model is more than sufficient for various applications as illustrated in the next section.

6 Applications

In the previous sections we showed that the use of visual significance results in smoothing that better correlates perceived strength of edges. However, applications like image abstraction through edge preserving smoothing or detail enhancement produce images whose quality is judged aesthetically. Thus, despite the obvious differences between the perceptual and non-perceptual methods, one can not objectively prove that a visually significant edge model produces better results. In this section we present three applications that rely on importance of image features, and thus the improvement through a perceptual model can be demonstrated through examples. All results are generated using the extended EAW framework. The edge maps used in image retargeting and panorama stitching are generated by using the inverse of Equation 1 as discussed in Section 3.

6.1 Image Retargeting

Several techniques were recently proposed to allow content-aware image and video retargeting [Avidan and Shamir 2007; Wang et al. 2008; Rubinstein et al. 2009]. The central part of those approaches is usually an *importance map* (energy function) that describes the importance of areas in the image. Using the map, the retargeting operator then preserves the important areas at the expense of less-important ones. Several possibilities of the importance map construction were proposed [Avidan and Shamir 2007], however a simple Sobel operator was utilized in many cases.

The visually significant edges are a natural candidate to construct such importance map in a perceptually more convincing way. We

show the results of seam carving image resizing operator [Avidan and Shamir 2007] using traditional importance map and the new map calculated by our technique in Figures 9 and 10. The traditional technique removes more visually significant areas than when we build importance map using our method. Our results indicate that the difference between both methods is especially significant if the visually significant details are located in dark image regions. While the perception of brighter details ($> 100 \text{ cd/m}^2$) can be approximated by a simple compressive logarithmic function, our method has the advantage of faithfully modeling perception in all luminance levels and taking masking into account, and thus overall produces more reliable results (Fig. 10 (c) and (d)). In fact, the success of particular importance map construction varies with the input images and the absence of a universal retargeting operator led to the proposal of a hybrid approach combining several techniques [Rubinstein et al. 2009]. Our results suggest that visual significance can be guideline in importance map computation and can provide a basis for more sophisticated retargeting operators.

An advantage of our approach is that it allows perceptually based retargeting on not just ordinary, but also high dynamic range images. In images consisting of mostly bright regions ($> 100 \text{ cd/m}^2$) a simple logarithmic non-linearity may be sufficient to approximate the perception of luminance. However, this method is less precise in darker regions where Weber's law doesn't hold (compare Fig. 10 (f) and (g)). Moreover, visual masking may have a significant effect in images containing many details (Fig. 9).

That said, we found that first producing a tone mapped "dual" image, and then performing the retargeting on the original HDR image using the edge strengths computed on the dual image to work well in some cases. However, the type of tone mapping operator and suitable parameter setting is an open question, and requires manual interaction in comparison to our fully automated method.

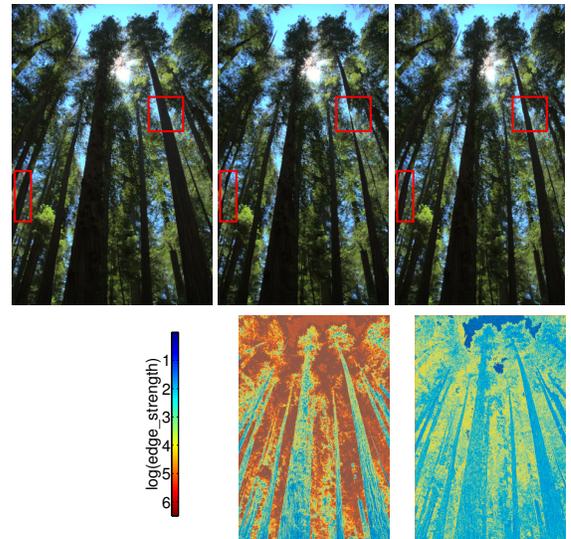


Figure 9: HDR image shrinking by seam carving (150 pixels horizontally). First row left: original HDR image. Middle: result when the Sobel operator is used for importance map construction. Right: result using the proposed visually significant edges. Images are tone mapped [Drago et al. 2003] for the display purposes. Second row: edge strength maps. Left: edges detected by Sobel operator in the input HDR image. Right: visually significant edges – note the differences in absolute values and in the ratios of edge strengths (due to the JND scaling), and the structural differences in the edge map (due to the masking).

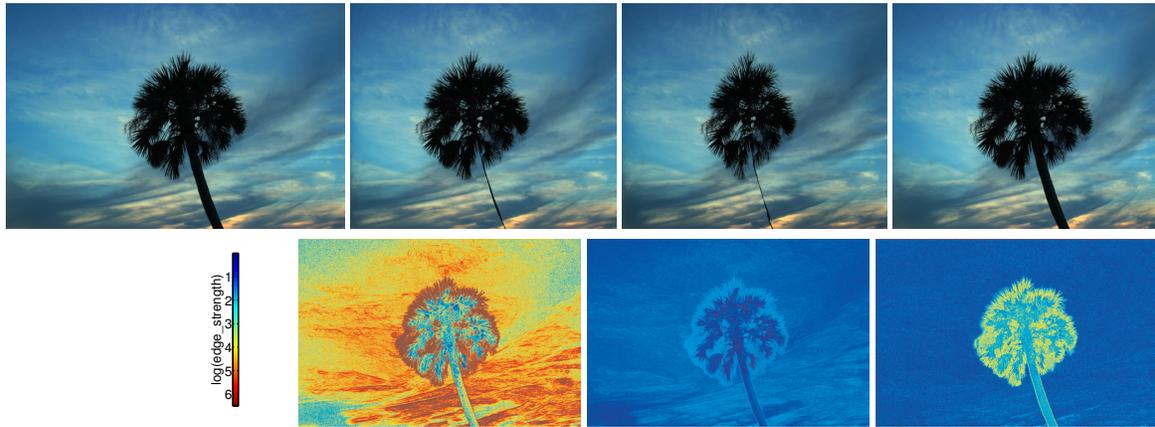


Figure 10: HDR image shrinking (400 pixels horizontally) by seam carving. First row: (a) original HDR image, (b) Sobel operator overestimates the strength of edges in the sky, which results in carving of the visually important palm tree, (c) results are similar if the Sobel operator results are compressed by the logarithm function, (d) the proposed method results in less distorted image appearance, especially evident at the tree’s body. Images are tone mapped [Drago et al. 2003] for the display purposes. Second row: (e,f,g) corresponding edge strength maps.

6.2 HDR Tone Mapping

As mentioned in experimental evaluations [Kuang et al. 2007; Čadík et al. 2008], the goal of tone mapping is manifold: some tone mapping operators are focused on compressing the image luminance while preserving the overall scene appearance. For example, the outcome of such an operator applied to a dark scene would not reproduce the details that are not visible by the human eye due to insufficient lighting. The other group of tone mapping operators on the other hand focuses on preserving as many scene details as possible irrespective of their visibility magnitude.

The tone mapping from the original edge avoiding framework [Fattal 2009] can be classified as strictly detail preserving. In the spirit of previous decomposition-based approaches [Tumblin and Turk 1999; Fattal et al. 2002; Durand and Dorsey 2002; Farbman et al. 2008], the technique flattens the coarsest scale of the EAW image decomposition by factor β and the other scales are progressively compressed so that the wavelet coefficients in a coarser scale are decreased more than in a finer scale (by factor γ^k , where k is the scale). This corresponds to an observation that the coarser scales often contain very high magnitude differences and should be therefore compressed much more than the finer scales (details) that we usually aim to preserve. The technique operates on *logarithm* of the input luminance that can be thought of as a simple approximation of human luminance perception, but having not accounted for other prominent perceptual phenomena (e.g. the perception of contrast), the results look unnatural, see Fig. 11 (left).

The results produced by the technique mentioned above may be suitable for certain scenarios (e.g. the best reproduction of details), but not for reproducing the appearance of a scene. However, we can achieve much better results (in this sense) by replacing the logarithm function with the perceptual framework proposed in this paper. We thus obtain image decomposition coefficients that are closer to the human visual system response (accounting for phenomena described in Section 4) and those are then compressed in a same way as above for the display purpose. As expected, the results are then more natural renditions of the original HDR images and preserve the scene appearance, see Fig. 11 (right).

6.3 Panorama Stitching

An HDR panorama generation approach proposed by Ward [2006] makes use of edge maps to stitch adjacent images of a scene. In this method images are decomposed into two layers: a low pass layer that corresponds to $1/16^{th}$ of the image’s original resolution and a high frequency layer. The low frequency layers of adjacent images are blended together using a sinusoidal weighting function, whereas the high frequencies are spliced at locations containing strong edges. The method is guided by a compound edge map E obtained as a combination of edge maps of pairs of overlapping images (E_{left}, E_{right}). We adopted the following technique to construct the compound edge map:

$$E = \max(E_{left} \cdot E_{right}, 0). \quad (5)$$

In other words: if there is a strong edge in the left image, but not in the right image, then this is possibly due to a misalignment and should not be preferred for splicing. On the other hand, locations containing strong edges with the same sign in both images are strong candidates for splicing.

For panorama stitching application, we inverted the neighborhood masking in our model, so that it amplified the masked edges. This is motivated by observation that the masked edges also mask the seams so that they are less disturbing in the final panorama. We empirically found that multiplying R with $(2 \cdot Neighborhood_masking)^2$ to work well in practice. We compare the results obtained using our technique and the traditional Sobel operator in Fig. 12. The source images were inverse tone mapped prior to processing by simple contrast stretching.

7 Conclusion

We presented a method that localizes image edges and scales their strength proportionally to their visual significance. We discussed a simple and efficient HVS model that accounts for prominent features of the visual system such as luminance adaptation, spatial frequency sensitivity and visual masking. In our experience the visual significance computation in EAW framework increases the computation time by 30 – 50%.

The HVS model is integrated into the edge avoiding wavelet frame-

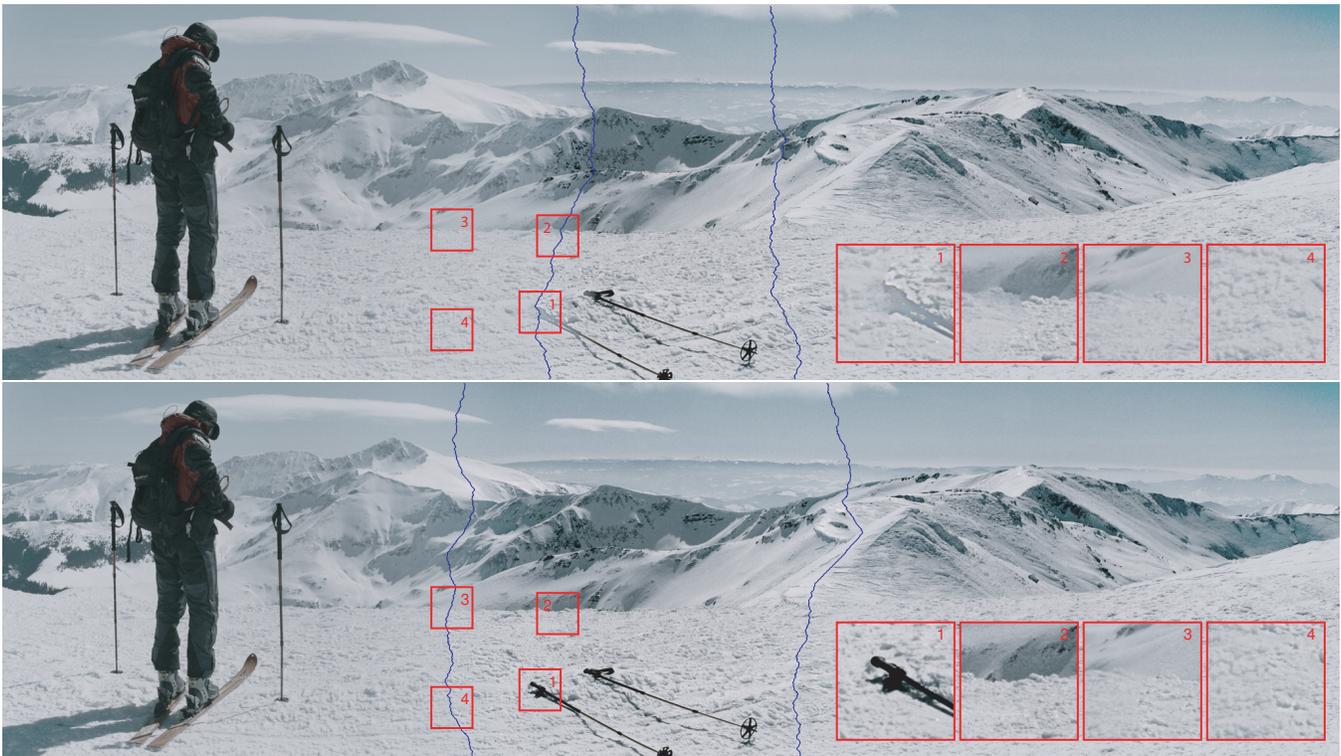


Figure 12: An HDR panorama stitched from three different, not precisely aligned pictures using Ward's technique [Ward 2006]. Top: the result obtained using Sobel operator; Bottom: the result using the proposed visually significant edges. The images are tone mapped [Reinhard et al. 2002] for display purposes.

work which provides a convenient basis for edge preserving image decomposition, and also extraction of edges by inverting the edge-stopping criterion. The choice of the framework is not crucial for specialized applications that rely either solely on image decomposition or edge extraction. For example, the HVS model can be applied to multi-scale image gradients for the former type of applications, or to an image pyramid obtained through bilateral filtering for the latter type of applications. The wavelet framework is convenient in the sense that it can serve both purposes in one framework, and is faster than others in decomposition.

The main limitation of this work is the absence of models for higher level mechanisms of the visual system such as gestalt properties and prior knowledge. Unfortunately modeling those mechanisms is not trivial because of their complexity and consequently the hardness of designing reproducible experimental setups to determine their effects.

In the light of recent work [Cole et al. 2008] that shows luminance edges are in fact prominent image features, we believe that the visually significant edges are good candidates for determining the richness of detail in images. Such a measure, combined with others such as image brightness, overall contrast and colorfulness can provide a good estimate of image quality in the absence of a reference image (no-reference image quality assessment). As a future direction we would like to investigate the possibility of designing such a metric that utilizes visually significant edges.

References

- AVIDAN, S., AND SHAMIR, A. 2007. Seam carving for content-aware image resizing. In *ACM SIGGRAPH*, ACM, New York, NY, USA.
- AYDIN, T. O., MANTIUK, R., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2008. Dynamic range independent image quality assessment. In *ACM SIGGRAPH*, ACM, 1–10.
- BARTEN, P. G. 1999. *Contrast sensitivity of the human eye and its effects on image quality*. SPIE – The International Society for Optical Engineering.
- BOLIN, M. R., AND MEYER, G. W. 1998. A perceptually based adaptive sampling algorithm. In *Proc. of ACM SIGGRAPH 1998*, 299–309.
- ČADÍK, M., WIMMER, M., NEUMANN, L., AND ARTUSI, A. 2008. Evaluation of HDR tone mapping methods using essential perceptual attributes. *Computers & Graphics* 32, 330–349.
- CANNY, J. 1986. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 8, 6, 679–698.
- COLE, F., GOLOVINSKIY, A., LIMPAEGER, A., BARROS, H. S., FINKELSTEIN, A., FUNKHOUSER, T., AND RUSINKIEWICZ, S. 2008. Where do people draw lines? *ACM Transactions on Graphics (Proc. SIGGRAPH)* 27, 3 (Aug.).
- DALY, S. 1993. The visible differences predictor: An algorithm for the assessment of image fidelity. In *Digital Images and Human Vision*, A. B. Watson, Ed. MIT Press, 179–206.
- DECARLO, D., AND SANTELLA, A. 2002. Stylization and abstraction of photographs. *ACM Transactions on Graphics* 21, 3 (July), 769–776.
- DRAGO, F., MYSZKOWSKI, K., ANNEN, T., AND N.CHIBA. 2003. Adaptive logarithmic mapping for displaying high contrast scenes. *Computer Graphics Forum* 22, 3.
- DURAND, F., AND DORSEY, J. 2002. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Transactions on Graphics (Proc. SIGGRAPH)* 21, 3, 257–266.
- ELDER, J. H., AND GOLDBERG, R. M. 2001. Image editing in the contour domain. *IEEE Transactions on Pattern Analysis and*



Figure 11: HDR image tone mapping without (left column) and with our HVS model (right column). The original method [Fattal 2009] preserves as many image details as possible at the cost of overall scene appearance. Our method is more balanced in terms of reproduction of scene appearance and detail preservation.

- Machine Intelligence* 23, 3, 291–296.
- FARBMAN, Z., FATTAL, R., LISCHINSKI, D., AND SZELISKI, R. 2008. Edge-preserving decompositions for multi-scale tone and detail manipulation. In *SIGGRAPH '08: ACM SIGGRAPH 2008 papers*, ACM, New York, NY, USA, 1–10.
- FATTAL, R., LISCHINSKI, D., AND WERMAN, M. 2002. Gradient domain high dynamic range compression. *ACM Transactions on Graphics (Proc. SIGGRAPH)* 21, 3, 249–256.
- FATTAL, R. 2009. Edge-avoiding wavelets and their applications. *ACM Trans. Graph.* 28, 3, 1–10.
- FERWERDA, J. A., PATTANAİK, S. N., SHIRLEY, P. S., AND GREENBERG, D. P. 1997. A model of visual masking for computer graphics. In *Proceedings of SIGGRAPH 97*, Computer Graphics Proceedings, Annual Conference Series, 143–152.
- FREEMAN, W. T., AND ADELSON, E. H. 1991. The design and use of steerable filters. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 13, 9, 891–906.
- GEORGESON, M. A., MAY, K. A., FREEMAN, T. C., AND HESSE, G. S. 2007. From filters to features: scale-space analysis of edge and blur coding in human vision. *Journal of Vision* 7, 13.
- JANSEN, M. H., AND OONINCX, P. J. 2005. *Second Generation Wavelets and Applications*. Springer.
- KUANG, J., YAMAGUCHI, H., LIU, C., JOHNSON, G. M., AND FAIRCHILD, M. D. 2007. Evaluating HDR rendering algorithms. *ACM Transactions on Applied Perception* 4, 2, 9.
- LEGGÉ, G., AND FOLEY, J. 1980. Contrast masking in human vision. *Journal of the Optical Society of America* 70, 12 (Dec.), 1458–1471.
- LINDBERG, T. 1996. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision* 30, 2, 77–116.
- MANTIUK, R., MYSZKOWSKI, K., AND SEIDEL, H. P. 2006. A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Appl. Percept.* 3, 3, 286–308.
- MANTIUK, R., DALY, S., AND KEROFISKY, L. 2008. Display adaptive tone mapping. *ACM Transactions on Graphics (Proc. SIGGRAPH)* 27, 3, Article 68.
- MARR, D., AND HILDRETH, E. 1980. Theory of edge detection. *Proc. Royal Soc. Lond. B*, 207, 187–217.
- ORZAN, A., BOUSSEAU, A., BARLA, P., AND THOLLOT, J. 2007. Structure-preserving manipulation of photographs. In *Int. Symposium on Non-Photorealistic Animation and Rendering*.
- PATTANAİK, S. N., FERWERDA, J. A., FAIRCHILD, M., AND GREENBERG, D. P. 1998. A multiscale model of adaptation and spatial vision for realistic image display. In *SIGGRAPH '98 Proceedings*, 287–298.
- PELLEGRINO, F. A., VANZELLA, W., AND TORRE, V. 2004. Edge detection revisited. *Systems, Man, and Cybernetics, Part B, IEEE Transactions on* 34, 3, 1500–1518.
- PEREZ, P., GANGNET, M., AND BLAKE, A. 2003. Poisson image editing. *ACM Transactions on Graphics* 22, 3, 313–318.
- PERONA, P., AND MALIK, J. 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, 629–639.
- REINHARD, E., STARK, M., SHIRLEY, P., AND FERWERDA, J. 2002. Photographic tone reproduction for digital images. In *SIGGRAPH '02*, ACM Press, 267–276.
- RUBINSTEIN, M., SHAMIR, A., AND AVIDAN, S. 2009. Multi-operator media retargeting. In *ACM SIGGRAPH*, ACM, New York, NY, USA, 1–11.
- RUZON, M. A., AND TOMASI, C. 1999. Color edge detection with the compass operator. In *Computer Vision and Pattern Recognition*, vol. 2, 166 Vol. 2.
- SWELDENS, W. 1997. The lifting scheme: A construction of second generation wavelets. *SIAM J. Math. Anal.* 29, 2, 511–546.
- TAYLOR, M. M., AND CREELMAN, C. D. 1967. PEST: Efficient estimates on probability functions. *The Journal of the Acoustical Society of America* 41, 4A, 782–787.
- TUMBLIN, J., AND TURK, G. 1999. LCIS: A boundary hierarchy for detail-preserving contrast reduction. *Proceedings of SIGGRAPH*, 83–90.
- UYTTERHOEVEN, G., ROOSE, D., AND BULTHEEL, A., 1997. Wavelet transforms using the lifting scheme.
- WANDELL, B. 1995. *Foundations of Vision*. Sinauer Associates.
- WANG, Y.-S., TAI, C.-L., SORKINE, O., AND LEE, T.-Y. 2008. Optimized scale-and-stretch for image resizing. In *ACM SIGGRAPH Asia*, ACM, New York, NY, USA, 1–8.
- WARD, G. 2006. Hiding seams in high dynamic range panoramas. In *APGV '06: Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization*, ACM, 150–150.
- WATSON, A. B., AND SOLOMON, J. A. 1997. A model of visual contrast gain control and pattern masking. *J. Opt. Soc. Am. A*, 14, 2379–2391.
- WHITTLE, P. 1986. Increments and decrements: Luminance discrimination. *Vision Research* 26, 10, 1677–1691.
- WILSON, H. 1980. A transducer function for threshold and suprathreshold human vision. *Biological Cybernetics* 38, 171–178.
- ZENG, W., DALY, S., AND LEI, S. 2000. Visual optimization tools in JPEG 2000. In *Proc. of Inter. Conf. on Image Processing*, vol. 2, 37–40 vol.2.
- ZIOU, D., AND TABBONE, S. 1997. Edge detection techniques - an overview. Tech. rep., International Journal of Pattern Recognition and Image Analysis.