24.1: Measuring the Perceived Contrast of Natural Images

Andrew M. Haun and Eli Peli
Schepens Eye Research Institute, Boston MA
Department of Ophthalmology, Harvard Medical School, Boston MA

Abstract:

We show how perceived contrast of natural images, as a function of spatial frequency, is measurable using a novel form of the psychophysical 'classification image' paradigm. Contrast spatial frequency weighting functions peak at mid-high frequencies, and are more stable than predicted by common definitions of image contrast or contrast sensitivity.

1. Objective and background

Perceived contrast is well defined for narrowband spatial patterns. However, perceived contrast of complex, broadband patterns – namely of photographic imagery – has not been subjected to a similar degree of study, and in fact there is no empirically-based definition of the perceived contrast of a complex image. To date, there is no answer to the question, "how does a human observer evaluate the contrast of a complex image?"

Peli [1] pointed out that the multi-scale nature of contrast processing in the human visual system would require at least a multi-scale description of image contrast, i.e. a description at multiple spatial frequencies (s.f.'s), if the description is to have any relevance to human perception. However, he addressed only the threshold aspect of contrast perception and assumed that suprathreshold contrast perception follows the contrast constancy of Georgeson and Sullivan [2]. An empirically-based means of estimating the perceptual impact of an image's contrast, or lack thereof, would be particularly useful in generating meaningful simulations of human image perception, could contribute to measures of perceived blur or sharpness, and ultimately provide a means of measuring the perceived quality of a broadband image.

Here we present a method for measuring the perceived contrast of narrowband spatial components as they constitute broadband imagery, in the context of other components which are measured simultaneously. These weighting functions vary predictably with certain image statistics, and results can be simulated effectively with a modified model of narrowband perceived contrast.

2. Methods

The method is a variation on the 'classification image' technique used to reveal perceptual templates used by observers in visual identification tasks [3, 4]. On each trial, an image is divided into seven 1-octave (raised cosine) spatial frequency bands, and an 8th band which contains the high spatial frequency residual. The amplitude of each band is increased or decreased by a random amount, with the constraints that one band is increased by the maximum of 0.4 log units, and that another band is decreased by the maximum of 0.4 log units. This procedure is repeated twice, and the two sets of randomly re-weighted spatial frequency bands are reassembled into two new images. The new images are visibly distorted but are still clearly recognizable (Figure 1).

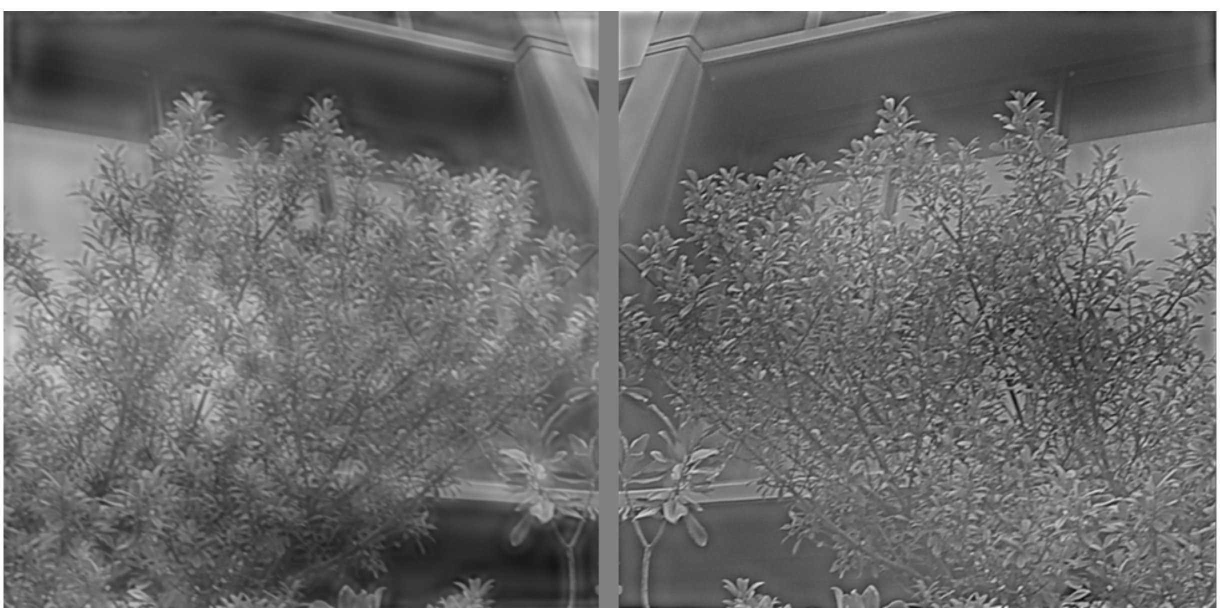

Figure 1. Example of a mirror-image stimulus pair used on one trial of the experiment. Spatial frequency band contrasts are randomly increased or decreased for one image; the same random weights are then applied in a different order to the second image. So, both images have the same amount of contrast perturbation, but applied to different spatial frequencies. Subjects are instructed to choose the image which has the greater overall contrast, defined as "brightest brights, darkest darks, of any and all sizes, over the entire image area". Subjects are instructed to not let single features drive their decisions, and to take into account the whole area of each image.

 ISSN 0097-966X/11/4201-0302-$1.00

The source images were a set of 516 scene photographs used in previous studies of natural scene statistics [5], displayed at 480x480px on a CRT display (Sony Trinitron p1103) set at 768x1024 resolution, and viewed at a distance of 1m, where each subtended 10.3°va. The two images are then displayed side-by-side, untimed, and an observer is tasked with choosing which image has the higher contrast.

The observer's instructions are to compare the apparent intensity of light and dark regions between the two images, and to choose the image with darker dark regions, lighter light regions, and the overall larger apparent range of luminances. Subjects are specifically instructed to avoid using single features to make their decisions, and to estimate contrast by taking into account the whole area of each image. This procedure is repeated for 500 different images in one block of sequential trials. The task is repeated 4 times (same images, different random weights), for a total of 2000 trials per observer.

At the end of the experiment, the strength of the relationship between trial-to-trial s.f. weights and observer decision (i.e., an image is chosen *and* rejected on each trial) is calculated as a vector of regression coefficients m_i relating the observer's decision d (chosen = 1, rejected = 0) to the contrast weights w_i (Equation 1):

$$d = \varepsilon + \sum_{i=1}^{8} m_i w_i \qquad (1)$$

The resulting coefficients function (Figure 2) shows how different spatial frequencies figure into the observer's judgments of image contrast. Thus we describe these as *weighting functions*, although in this context they are weights determining the observer's decision in the task. So, the functions can be taken only to describe the relative weighting of different components in judgments of perceived contrast, *not* the absolute perceived contrast of the images.

3. Results

Resulting weighting functions are peaked at mid-high spatial frequencies (between 1.5 and 6 cpd), with the precise function shape varying across different observers (Figure 2). Weight magnitudes, signifying the strength of the relationship between observer decision and spatial frequency, were very similar across observers. For all observers, low s.f.s were negatively weighted, meaning that increased low s.f. contrast made it unlikely that an image would be seen as having 'higher contrast'. Likewise, the high s.f. residual band was always negatively weighted, perhaps because it contained imaging noise and details irrelevant to the content of the test images.

The weighting functions are also found to be dependent on image statistics. For example, for Figure 3 the set of 516 trial images were binned into 10 sets according to their edge density [6] (the mean of the Canny-filtered original images, using the default settings of the Matlab 7.5 *edge* function, with higher values corresponding to images with denser textures and less undetailed area). Contrast weights were computed from trial data for each image set. Increasing edge density corresponds to a more peaked weighting function (red denoting positive, blue negative weight strength) and a progressively higher spatial frequency peak.

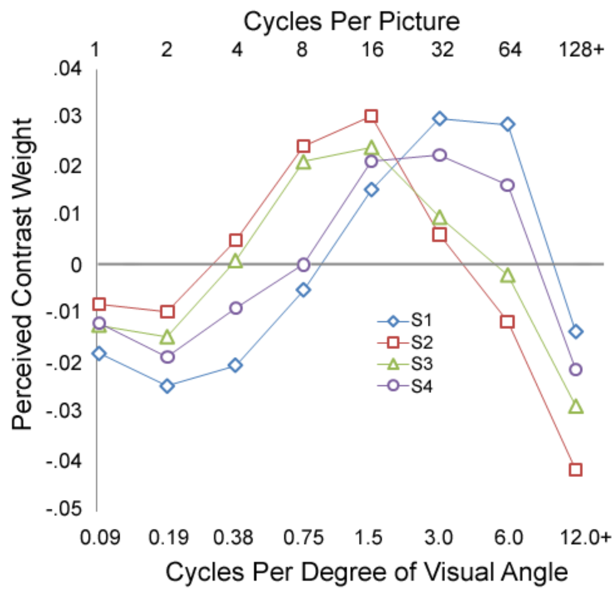

Figure 2. Regression coefficients (perceived contrast weights) for spatial frequency bands for four observers' judgments of broadband image contrast. Coefficients describe the strength (magnitude) and direction (positive or negative) of the relationship between changes in s.f. band contrast and observer choice. Positive coefficients indicate that increasing band contrast increased the likelihood of an image being chosen; negative coefficients indicate that decreasing contrast increased the likelihood. All observers show contrast weighting function peaks at mid-high spatial frequencies, at or above the peak of the normal contrast sensitivity function.

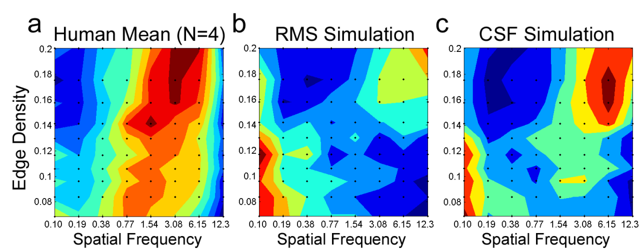

Figure 3. Interpolated weighting maps for human (n=4) and simulated observers, showing weighting relationships between spatial frequency and edge density. All observers, human and simulated, were run through 2000 trials of the experiment. X-axis is spatial frequency in cycles per degree, Y-axis is Canny edge density. Black dots represent bin locations, the vertices of the maps. a) Human observers consistently show a peak in spatial frequency weighting between 1.5 and 3 cpd, with changes in peakedness and a gradual shift in peak with image type. Simulated observers do not show such stable behavior – RMS (b) and CSF observers (c) are chiefly dependent on low s.f. contrast for edge-sparse images, and transition abruptly to high s.f. contrast when images are more edge dense.

Different definitions of image contrast can be tested by building simulated observers and running them through the same experiment as the human observers. From trial to trial, images are generated and processed according to the specified model, which then chooses one image based on a set definition of perceived contrast. A simulation that chooses images based only on global RMS contrast, for example, fails completely to produce peaked weighting functions, instead relying primarily on low s.f. contrasts for most image types (Figure 3b). Contrast sensitivity function (CSF)-based simulation [7], which varies in its sensitivity to contrast with spatial frequency, comes closer, but does not produce stable peaked weighting functions when set with, e.g., 'standard observer' parameters [8], instead displaying an abrupt transition between low-pass and high s.f.-peaked weighting functions as edge density increases (Figure 3c). We suggest that the high s.f. perceived contrast peak of human observers is due to low-s.f.-biased contrast gain control [6,9], although the relevant parameters require further empirical investigation.

4. Discussion

These results demonstrate that the perceptual impact of an image, and the way its contrast is interpreted by an observer, is dependent on the structure of the image, thereby suggesting that perceived contrast of complex imagery is not an entirely passive process. Individual differences in the weighting functions also may correspond with previously described differences in preference for video enhancement by normal observers. Satgunam *et al.* [10] have shown that some observers perceive relatively smooth, unenhanced video as having the most pleasing appearance, while others prefer much more strident, sharpened, enhanced video. We believe that these groups may be weighting contrast differently, with the smooth-preference group having relatively higher high-frequency weights, thereby making high-s.f. enhancement excessive, while observers with lower s.f. contrast weighting peaks are better able to appreciate the additional content revealed by video enhancement algorithms.

5. Impact

The importance of this study lies in the ability of this method to describe in psychophysically valid terms the contrast appearance of a photographic image. Contrast is a primary component of most image quality metrics, and drives virtually all models of early visual processing. Most of what is known of the perceived strength of a given contrast pattern, however, is drawn from studies using simple patterns, and it is demonstrated here that these studies are not sufficient to predict the form of the perceived contrast weighting function. We expect that with the ability to more precisely measure and simulate perceived contrast of broadband imagery, models of perceived quality should be significantly improved.

6. Acknowledgements

Supported in part by NIH grants EY05957, EY16093, and a grant from Analog Devices Inc.

7. References

[1] E. Peli, "Contrast in complex images," Journal of the Optical Society of America - A **7(10)**, 2032-2040 (1990).

[2] M.A. Georgeson, G.D. Sullivan, "Contrast constancy – deblurring in human vision by spatial frequency channels," Journal of Physiology **252**, 627-656 (1975).

[3] C.K. Abbey, M.P. Eckstein, "Classification image analysis: Estimation and statistical inference for two-alternative forced-choice experiments," Journal of Vision **2(1)**, 66-78 (2002).

[4] D.M. Levi, S.A. Klein, I.N. Chen, "What is the signal in noise?" Vision Research **45**, 1835-1846 (2005).

[5] B.C. Hansen, E.A. Essock, "Influence of scale and orientation on the visual perception of natural scenes," Visual Cognition **12**, 1199-1234 (2005).

[6] P.J. Bex, S.G. Solomon, S.C. Dakin, "Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure," Journal of Vision **9(10):1**, 1-19 (2009).

[7] M.W. Cannon, "A multiple spatial filter model for suprathreshold contrast perception", in Vision Models for Target Detection and Recognition, ed. Eli Peli (NJ, 1995), pp 88-117.

[8] A.B. Watson, A.J. Ahumada, "A standard model for foveal detection of spatial contrast," Journal of Vision **5**, 717-740 (2005).

[9] A.M. Haun, E.A. Essock, Contrast sensitivity for oriented patterns in 1/f noise: contrast response and the horizontal effect. Journal of Vision **10(10):1**, 1-21 (2010).

[10] P. Satgunam, R.L. Woods, P.M. Bronstad, E. Peli, "Factors affecting image quality preferences," SID Symposium Digest **41**, 94-97 (2010).