Automatic assessment and reduction of noise using edge pattern analysis in non-linear image enhancement

Daniel J. Jobson[‡], Zia-ur Rahman[†], Glenn A. Woodell[‡], Glenn D. Hines[‡]

[‡]NASA Langley Research Center, Hampton, Virginia 23681.
[†]College of William & Mary, Department of Applied Science, Williamsburg, VA 23187.

ABSTRACT

Noise is the primary visibility limit in the process of non-linear image enhancement, and is no longer a statistically stable additive noise in the post-enhancement image. Therefore novel approaches are needed to both assess and reduce spatially variable noise at this stage in overall image processing. Here we will examine the use of edge pattern analysis both for automatic assessment of spatially variable noise and as a foundation for new noise reduction methods.

1. INTRODUCTION

We have previously defined that the visibility limit in non-linear image enhancement[1,2] was set by image sensor noise.[3–5] This image enhancement process was based upon the use of visual measures for contrast, lightness, and sharpness[4] in a smart visual servo feedback computation.[3] Experiments with the visual servo suggested the need to add an additional assessment/enhancement control to improve performance in the presence of visually significant noise in the enhanced image. While noise can easily be determined by capturing dark image frames and calibrated to a range of exposure settings, this is not possible in general purpose imaging, where noise levels must be determined from the image itself. And, indeed, noise is a visually significant factor only when *feature* signal-to-noise ratios (SNRs) drop to low levels that approach unity. The *feature* SNR is distinct from the commonly used *root-mean-square* (RMS) SNR in that the feature SNR is not measured globally over the whole image, but rather regionally across the feature under consideration. Given that the application of powerful non-linear enhancement algorithms produces a highly spatially variable level of noise, sophisticated new methods are needed to both assess and reduce noise in the post-enhancement image. Further for our case, noise management methods should be integrated into the overall smart enhancement processing scheme as an advance toward an even more comprehensive approach to the improved visual representation in the digital image. The management of noise can be viewed as two cases which are not completely distinct but which rather merge gracefully. These two cases are: (1) moderate levels of noise can become visually significant in the enhanced darker regions or turbid regions of an image; and (2) high noise levels may completely mask all visual information in strongly enhanced regions of the image.

In this paper, we will re-examine the fundamental visual nature of noise in images and its meaning in terms of spatial resolution, examine some preliminary concepts for determining spatially variable noise by using edge pattern analysis, and explore some approaches to noise reduction for the moderate noise case, and extracting completely masked features from the high noise case. For the former, we find that noise is essentially a reduction in spatial resolution that is best accommodated by a variable spatial resolution by preferentially averaging (blurring) noisy regions. For the latter, we find that spatial averaging does not extend sufficiently to encompass this situation since the amount of averaging required leaves virtually no spatial resolution. Therefore an entirely different approach is required for the high noise case, one that employs large spatial support matched filtering. However this means that we are confronted by exactly the same problem as visual recognition itself: the combinatorial explosion of having to test too many image regions against too many matching templates does not allow a practical scheme. Visual recognition runs into a similar combinatorial explosion by having to extract near-infinite arbitrary patterns and compare them to a near infinite set of templates stored in memory.

Contact: DJJ: daniel.j.jobson@nasa.gov; ZR: zrahman@as.wm.edu; GAW: glenn.a.woodell@nasa.gov; GDH: glenn.d.hines@nasa.gov;

Figure 1. Illustration of non-linear image enhancement and impact of noise on visibility. (a) original image (b) visual servo enhancement (c) enhancement with moderate level of noise injected in original (d) enhancement with very high level of noise injected in original

Both of these processes could be called "infinity-squared" problems. Therefore a practical solution to the high noise case must necessarily find a way to whittle down the "infinity-squared" to some tractable set of limited patterns, yet still succeed for the arbitrary patterns that might be encountered in any image. Figure 1 illustrates two enhanced images that represent these two stages of noise.

2. EDGE DETECTION OF NOISE PATTERNS

We would like to construct an edge detection scheme that has, at least, some perceptual validity. Namely, we would like the detected edges for both noise and feature patterns to accord with the visual perception of these phenomena in real images. Noise that is not perceptible, and likewise, features that are not perceptible in the enhanced image would not be considered by the edge detection procedure. Neurophysiological evidence suggests that visual perception has some built-in noise evasiveness. Visual perception is often described as a two-scale process, sometimes classified as the parvocellular (p) and magnocellular (m) subsystems. This is both an anatomical distinction[6] and a physiological distinction.[7,8] In primate vision, the p-channel is associated with high acuity and coarse contrast sensitivity, while the m-channel is associated with coarse spatial resolution and high contrast sensitivity. This arrangement appears to eliminate all but large noise spikes from the p-channel, and provides some noise smoothing benefits due to spatial averaging in the m-channel.

We have devised a scheme that does, at least approximately, capture our visual perception of noise and visual features in complex natural images. Figure 2 shows some examples of the results of this two-scale edge detection to demonstrate this approximate perceptual validity. The edge operator that we use is the commonly used Difference-of-Gaussian (DOG) convolution with zero-crossing edge detection. As an analog for the p- and m-channels, we use two scales of this operator and apply it to each color band. The p-channel analog is a smallest scale DOG convolution operator that takes the form,

$$p(x,y) = 1 - c_1 e^{-(x^2+y^2)/\sigma_1^2} \tag{1}$$

where (x,y) specifies the pixel location in Cartesian coordinates, $\sigma_1^2 = 4$ pixels and c_1 is adjusted so that $\int p(x,y)\,dx\,dy = 0$, and which incorporates the optics blur function into the construction of the convolution operator. For the m-channel,

$$m(x,y) = e^{-(x^2+y^2)/\sigma_2^2} - c_2 e^{-(x^2+y^2)/\sigma_3^2} \tag{2}$$

where $\sigma_2 = 1.1$ pixels, $\sigma_3 = 10$ pixels, and c_2 is adjusted so that $\int m(x,y)\,dx\,dy = 0$. This operator has about a 3 pixel center diameter, but the surround for the m-channel is much larger than three times the positive center diameter. This is somewhat larger than might be expected, but produces better results in diverse perceptual comparisons. Within this system, a high threshold value for zero-crossing detection is employed in the p-channel (which responds only to very high contrast edges), and a much lower threshold is used in the m-channel (which provides a much high contrast sensitivity, while spatially averaging the noise). No attempt was made at this early stage of investigation to incorporate the well known temporal response differences between the p- and m-channels (p-channel responds to sustained edge patterns, while m-channel responds to transient edge phenomena).

With this foundation we can proceed to the experimental study of noise. The edge detection results for some enhanced noisy images are shown in Figure 3. Several characteristics of noise do show up as being distinct from the visual feature information. We also note that noise can be perceived as textural information. In fact, some textural edge patterns, such as very fine weave textures at or near the spatial resolution limit, mimic noise rather convincingly. Some candidate distinctive characteristics between noise and features are:

1. edge continuity or connectivity—features have it, noise doesn't,
2. edge spatial densities—noise has high and regionally uniform densities, while features usually possess lower and much more variable densities. Even such dense feature data as printed text have word spaces and line spaces that prevent uniform high densities from persisting across significant regions of image space,
3. edge orientations—features do not generally have random orientations, noise does.

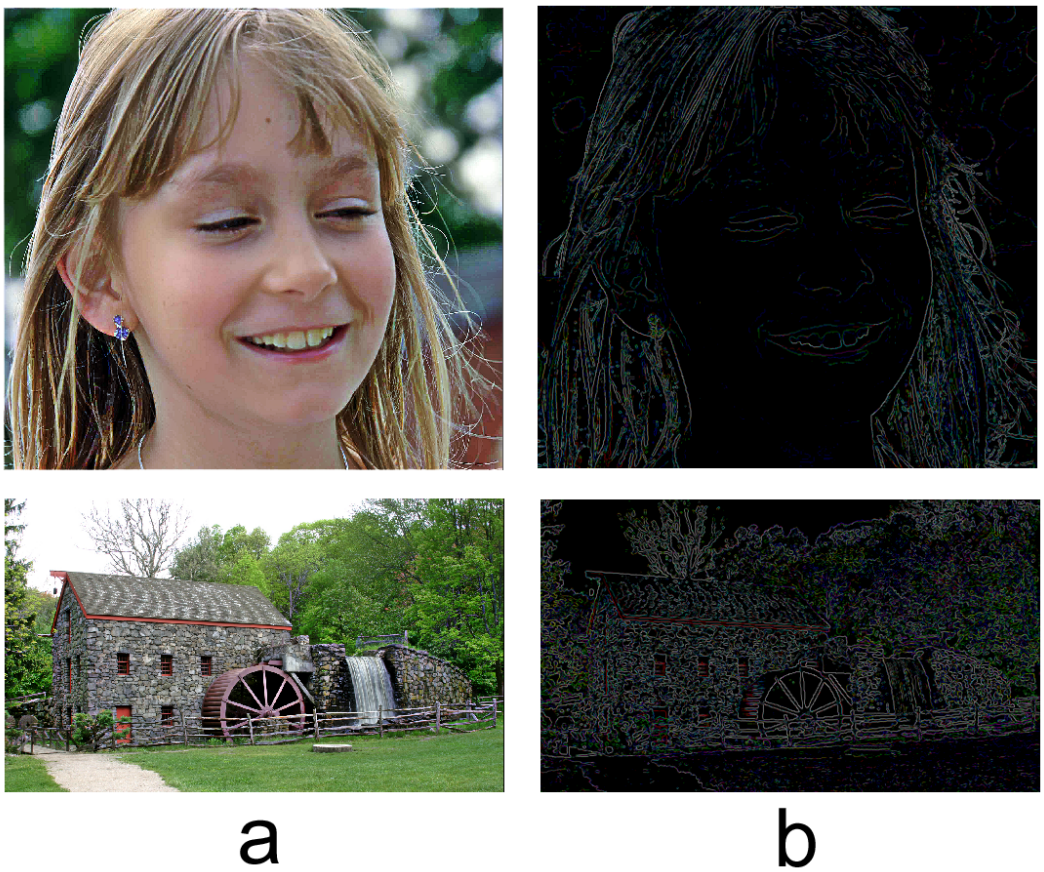

Figure 2. Two examples of post-enhancement edge detection (combined p- and m-channel results) to illustrate approximate perceptual validity of edge detection process (a) enhanced noisy image (b) detected edge patterns (c) detail of noisy image (d) blurred and sub-sampled version of (a) (e) edge pattern for (d)

One or more of these characteristics could form a sufficient basis for the quantitative assessment of spatially variable noise and, as such, be the foundation for remedial processing of noisy zones. We also examined the edge patterns for the much higher levels of noise. Again these same characteristics do show promise, especially the simple connectivity and high edge density. Noise produces uniform edge densities that approach the spatial limit set by the sampling by the edge detection operator.

Several hundred images of diverse types were tested for various levels of naturally occurring or simulated noise. Some preliminary comments on naturally occurring noise are needed to explain that real-world noise in current digital camera images is not as random as might be expected. Most current digital cameras undersample color and use color interpolation to "estimate" the full spatial color image. As a result noise patterns exhibit greater connectivity than would occur for truly random noise. Further, the frequent use of lossy JPEG image compression likewise has an impact on noise patterns and can impress a square block pattern on edge data that is associated with the block coding and color under-sampling of JPEG. The impact of these non-random influences is to establish a higher connectivity criteria for local edge patterns in any noise assessment classification. For current digital cameras, we can make a preliminary estimate of the connectivity criteria for naturally occurring noise at visible, but not obscuring, levels. For our case of two-scale edge detection, the connectivity criteria is that noise can exhibit 6–10 pixels of connectivity within each color band for the small scale detected edges, and 20–30 pixels of connectivity for the larger scale. This is, of course far above the connectivity exhibited by random noise which at a comparable moderate level produces connectivity of 3–6 pixels at the small scale, and 10–15 pixels for the large scale edge detection channel. Therefore the connectivity criteria for the simulated

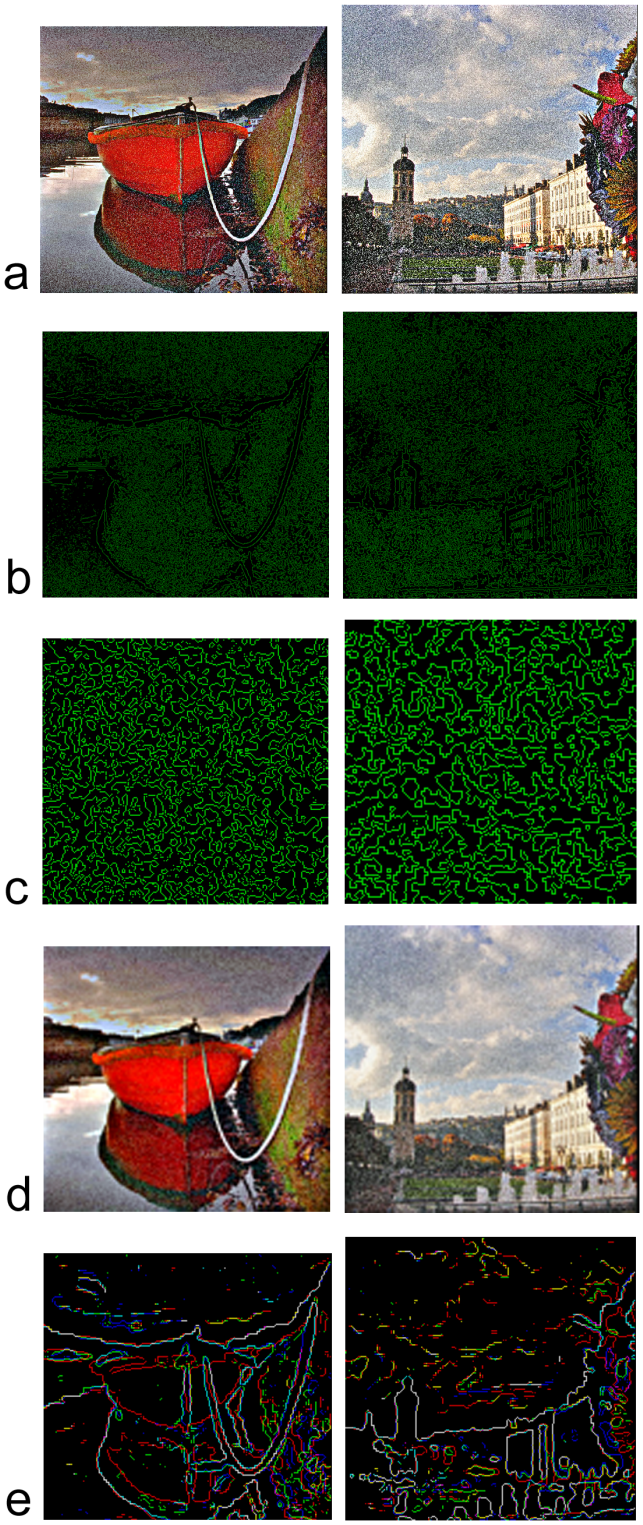

Figure 3. Two examples of the edge detection of noisy images (a) enhanced image—no noise (b) enhanced noisy image (c) vertical convolution image (d) horizontal convolution image

noise case has to be double that for the naturally occurring noise case.

For the case of very high noise levels which would be associated with enhanced images of very low light level or highly turbid scenes, connectivity becomes a more complicated issue for noise edge patterns. These high noise levels are associated with RMS SNR values that are higher than half of the image dynamic range. For the small scale edge detection most edges run together in a dense pattern of intersections—which, in itself, can serve as an indicator of noise. Thus, connectivity becomes an issue of *simple* connectivity, where each edge pixel can connect to only two nearest neighbors to form a single line segment. A pixel that does not have simple connectivity is considered as noise. The larger scale edge detection exhibits an entirely different phenomenology—very high levels of connectivity, with few complex intersections, are produced at very high and uniform edge density values. So for the m-channel edge pattern analysis, density and uniformity of density may be the required characteristic for correctly classifying a noise pattern, while simple connectivity may be adequate as a basis for the p-channel at high noise levels.

3. NOISE REMEDIATION AT MODERATE NOISE LEVELS

It is worthwhile to look ahead under the assumption that some of the previous noise assessment methods are workable, and examine approaches to noise remediation. If we pursue the idea that noise acts to lower spatial resolution, then we can proceed with remediation by replacing all noisy regions with low-pass filtered image values that retain the non-noise edge structures. This may not be entirely straightforward, since the retained edge structures may actually have to be represented as step edges that must be gracefully blended with the low-pass filtered intensity values. Considerable contrast information is available during the edge detection process. Therefore edge contrast could also be extracted at the same time as edges are detected. An alternate approach is simply to replace all regions classified as noise in the post-enhancement image with low pass values which should blend harmoniously with the regions containing feature edge structures. This latter approach is clearly the most straightforward. No doubt there are many other variations on replacing noisy regions with low pass image data.

4. FEATURE EXTRACTION FROM AT DEEP NOISE LEVELS

We can hardly be concerned with noise remediation for the extreme case where all of the image seems to be noise that, in a visual sense, completely masks any underlying structure in the scene. Of course, it would not be evident that there was any structure present, either visually or numerically, without substantial processing. In this case, the scene could be completely arbitrary and may just as easily be one containing no details—blank sky, for example—as it may be one containing a very complex scene such as a natural landscape. Remediation by low-pass filtering and replacement would not be possible because the amount of blurring required to separate noise from features would be so excessive that most real image data could be lost. The image, in essence, would be reduced to an uninformative point, or an exceedingly small thumbnail. Obviously, for this case, spatial averaging alone is insufficient. Something with greater precision is required to even attack this problem. By implication, this leads to fascinating questions, such as: "How much visual information can be recovered from the predominantly noise image?" We will not attempt to answer such questions in this paper.

We are, however, primarily interested in examining potential methods to extract any feature information, however crude, from the apparently noise-only image. We will assume that for the deep noise case, the connectivity criteria, edge orientations, and edge densities all fail to detect a significant number of features. From this starting point, the only potential approach seems to be one that would utilize some sort of matched filtering to extract features from noise. We can draw on psychophysical data that suggest that the eye-brain system is predisposed to vertical and horizontal structure in images. In fact we have applied vertical and horizontal oriented matched filters of a fairly small scale (ridge-valley operators with one pixel lobe widths, and seven pixel elongated extents) (Figure 4). These convolution operators can be thought of, when optics blur is accounted for, as representing small scale Gaussian-by-Difference-of-Gaussian forms that resemble the oriented "simple receptive fields" of the neurophysiology of the striate cortex. However, convolution images with these operators did NOT provide any significant ability to extract features at very high noise levels (Figure 5). Upon further thought we can see that this was to be expected. The matched filter still relies on noise averaging to achieve a

0.025	**0.150**	**0.375**	**0.500**	**0.375**	**0.150**	**0.025**
−0.025	**−0.150**	**−0.375**	**−0.500**	**−0.375**	**−0.150**	**−0.025**

Figure 4. Horizontal Gaussian by Difference of Gaussian convolution operator attempt at "matched filter". Note: Vertical operator is simply the horizontal operator rotated by 90°.

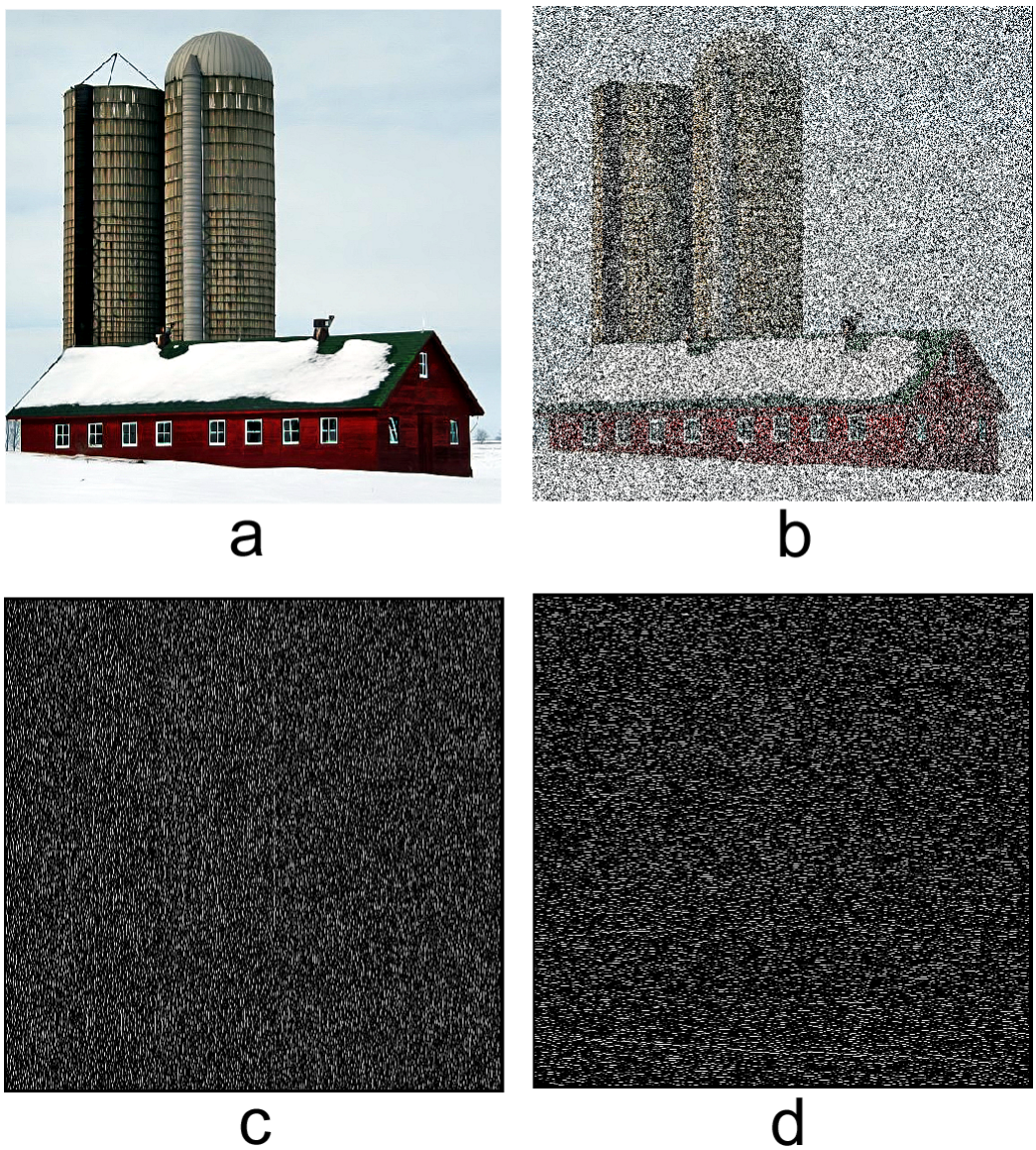

Figure 5. Examples of inability of vertical and horizontal small scale operators to extract features from high levels of noise. (a)vertical convolution image (b)horizontal convolution image

good level of performance, and our trial matched filters do not possess enough spatial support to substantially diminish noise.

Therefore, we are confronted at this point with the daunting similarity between this problem and that of general visual recognition: We need to be able to estimate a reasonably large spatial-scale matched filter without any a priori knowledge of the scene structure, or even whether the scene has any structure. This implies that at best we can only capture coarse, simple edges of considerable extent, and even this would require considerable computational burden. To capture fine structure, we would have to try a truly enormous range of matched filters that most likely will run into the implacable wall of combinatorial explosion. It may be possible to design an adaptive approach that could successfully maneuver between the trivial and the impossible. Here we can envision that for any very large spatial extent feature that is captured, a restricted set of fine structure variations of matched filters could be deployed. This is especially possible if visual classes can be defined that are restricted, yet still reasonably comprehensive. This line of thinking strikes us as being remarkable reminiscent of the successful recognition achieved by the immune system—namely molecular recognition of invading bacteria or viruses. Apparently the immune system begins its response to infectious disease by deploying a pre-set—coded into the DNA molecule—group of likely protein classes into the bloodstream. It then responds with a second stage attack whereby any "near-hits" from the initial deployment produce a large variational set from that class. This large variation set is deployed and is usually the successful stage of immune response whereby an exact "pattern" match is found, and this identification leads to marking the infectious agents for subsequent destruction by the white blood cells. While an analog for this immune response is not immediately obvious for visual information processing, this process may provide a point of departure for further investigation into this exceedingly challenging and intriguing problem.

5. CONCLUDING REMARKS

Since we are reporting concepts and results at an early stage of our investigation, it is appropriate to describe some remaining technical issues with the methods defined thus far for managing moderate and spatially variable noise, as well as providing some discussion of the future plans for the more imposing problem of extracting feature signals from deep noise. Several issues remain for the moderate noise case which include graceful merger of the p- and m- channel edge signals, and test filtering by edge connectivity constraints. Certain aspects of protocols for this are fairly obvious from the results thus far. These are:

1. p-channel detected edges should have priority over m-channel edges, and inhibit the larger scale channel's edges which are often spatially misaligned with the p-channel,
2. this inhibition of the larger scale channel edges cannot however extend over much regional space or it can inhibit desired edges not picked up by the p-channel,
3. the edge processing including connectivity filtering should take place within each color band and proceed from the small p-channel to the m-channel within each color band before final assembly of the two-scale color edge representation, and its noise filtered ultimate form.

With respect to the more monumental problem of extracting feature edges from deep noise, we can only speculate at this very early stage in this inquiry. The problem of finding features (almost) completely obscured by noise is clearly tantamount to that of general visual recognition. An effective practical solution seems to lie in the direction of establishing a workable grammar, or hierarchy, for visual structural pattern geometries. We need to determine general pattern classes—coarse, large-scale features—and then complete variations within each class—features corresponding to finer detail and smaller extents. Additionally, we need to determine restricted classes and their variations that encompass features that occur most commonly in the complex natural image. This immediately introduces a problem of needing to have precise structural specificity to address the fundamental problem with deploying these kinds of estimated matched filters: they must have sufficient spatial support to average noise in a dramatic manner. If this transition from general classes to specific matched filters can be achieved, then we can envision a practical scheme that is analogous to the molecular recognition processes of the immune system which successfully detects the specific 3-D molecular structures of pathological

bacteria and viruses. This line of thinking is really a statement about our hope that the power of comprehensive symbolic representation of features will be sufficient to conquer the combinatorial explosions that inherently block progress on visual recognition and its closely related problem, the extraction of weak feature signals from deep noise.

Acknowledgments

The authors wish the thank the Synthetic Vision Sensors element of the NASA Aviation Safety Program for the funding which made this work possible. In particular, Dr. Rahman's work was supported with the NASA cooperative agreement NCC-1-01030.

REFERENCES

1. D. J. Jobson, Z. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. on Image Processing* **6**, pp. 451–462, March 1997.
2. D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multi-scale Retinex for bridging the gap between color images and the human observation of scenes," *IEEE Transactions on Image Processing: Special Issue on Color Processing* **6**, pp. 965–976, July 1997.
3. D. Jobson, Z. Rahman, and G. A. Woodell, "Feature visibility limit in the enhancement of turbid images," in *Visual Information Processing XII*, Z. Rahman, R. A. Schowengerdt, and S. E. Reichenbach, eds., Proc. SPIE 5108, 2003.
4. D. Jobson, Z. Rahman, and G. A. Woodell, "The statistics of visual representation," in *Visual Information Processing XI*, Z. Rahman, R. A. Schowengerdt, and S. E. Reichenbach, eds., pp. 25–35, Proc. SPIE 4736, 2002. Invited paper.
5. Z. Rahman, D. Jobson, and G. A. Woodell, "Retinex processing for automatic image enhancement," in *Human Vision and Electronic Imaging VII*, B. E. Rogowitz and T. N. Pappas, eds., pp. 390–401, Proc. SPIE 4662, 2002.
6. Y. Shostak, Y. Ding, and V. A. Casagrande, "Neurochemical comparison of synaptic arrangements of parvocellular,magnocellular, and koniocellular geniculate pathways in owl monkey (aotus triviratus) visual cortex," *Journal of Comparative Neurology* **456**, pp. 12–28, 2003.
7. L. P. O'Keefe, J. B. Levitt, D. C. Kiper, R. M. Shapley, and J. A. Movshon, "Functional organization of owl monkey lateral geniculate nucleus and visual cortex," *Journal of Neurophysiology* **80**, pp. 594–609, August 1998.
8. N. H. Yabuta and E. M. Callaway, "Functional streams and local connections of layer4c neurons in primary visual cortex of the macaque monkey," *Journal of Neuroscience* **18**, pp. 9489–9499, November 1998.