8. CONCLUSION

The objective of this research was to test the hypothesis that a model of human vision could be used as the basis for making predictive measurements of the quality of medical images. This investigation required several distinct undertakings that were described methodically in the dissertation. First, the core model, a promising model for shape perception, was adopted and subsequently built upon to carry out two important visual tasks in medical images. Next, a basis for testing this novel approach was developed that consisted of the generation of a series of images that varied with respect to important physical parameters and a paradigm for collecting the human observer data. Finally, the model and human data were analyzed to render a verdict regarding the suitability of the model for further application as proposed. These steps represent a significant contribution to academic research; each of them, in the spirit of biomedical engineering, required exploration into diverse but complimentary disciplines.

The motivation for a computed measure of image quality is straightforward. Decisions must constantly be made about how to acquire or process the pictures of the body that allow visual diagnosis. These comparisons of modalities or machineries, or the decisions about how best to adjust physical or software parameters, can be done properly by the arduous process of experimentally measuring human performance. The gauge for quality in such experiments is task accuracy: when the reader determines the diagnosis most accurately the image is deemed “best.” The alternative to these experiments is to develop a computation that decides upon images that are maximally fit for human use.

This dissertation advocated and tested the notion that the best way to compute quality so that it can be predictive of human assessment is to explicitly model visual perception. A model of human vision might perform with the same abilities and disabilities as those of the human and in that way make quality determinations that parallel human assessment. The core model, which posits mechanisms for object representation, is well-suited for this approach: there is psychophysical evidence for its validity, and it can be used to perform shape-based medical image estimation tasks.

This research proposed a framework for testing this model-based approach. Two modalities, angiography and portal imaging, were chosen for exploration. Shape tasks that are performed with those systems were identified. Stenosis estimation, a judgment that quantifies vessel constriction, and treatment field clearance, a distance estimate that says how close the radiation is to the spinal cord, served as the representative tasks. Images were generated or simulated that varied with respect to physical variables describing the acquisition or processing method. The core model was modified appropriately so that the estimation information could be extracted following core analysis. Finally, human observer experiments were conducted to obtain the data against which the model was compared.

In the research, a sufficient but not necessary condition for the application of the model approach was tested: the differences between the model and human estimation errors had to be virtually constant throughout the parameter space. It is evident from the graphics and reinforced from the statistics that this was not achieved. In most cases the model behavior made sense in light of the expectations for the effects of the parameters. Angiographic noise and blur did, in a general sense, degrade the accuracy of the model. Higher settings of the SHAHE contrast parameter, something that was expected to induce an acknowledged artifact at the field edge, did cause systematic increases in model estimates. Yet the model’s tracking of the human data was not precise enough. The model’s performance was not such that it could be trusted, with similar images under similar conditions, to land upon those parameter settings that maximized human accuracy.

The estimation models performed on the whole remarkably well. Model percent stenosis estimates were rarely more than five percent narrower or wider than the true vessel constriction. The model’s deviation from human accuracy for the stenosis estimation task was normally less than five percent as well. The standardized differences showed that the model accuracy was typically within a single standard deviation of the human mean. For the distance estimation task, the ten percent differences in estimation accuracy are small given the complexity of the task. The model produced very clear and understandable trends as a function of the SHAHE parameters.

There were several reasons why the model-based predictions might not have been entirely consistent with the human data. The medialness kernel for the vessel estimation task was chosen to be the Laplacian of a Gaussian, a robust operator in the presence of blur and noise but probably not one that is analogous to any mechanism in the visual system. Second, ridge initiation and following often occurred with poor success in noisy angiograms or portal images. Many improvements to the ridge algorithms have been proposed and implemented since the time of this work that might improve these measurements from medialness that produce the core. More importantly, the way that the task was carried out once a core was obtained was a likely source of discrepancies. Stenosis estimates were computed from a vessel core by sliding a window along it and determining special positions at which to take width estimates. Gap object width estimates came from the mean in a positional region of the core. Those implementation decisions were made sensibly, but without theoretical foundation. Finally, there was some question, particularly in the portal imaging study, whether the human observers performed the task consistently and were able to use the measuring tools to indicate precisely their estimates.

The contributions that this work represent are several. What was proposed is a method for evaluating medical image quality with respect to an estimation task. There are few other measures in that regard, and yet estimation tasks are prevalent in the radiologist's interpretation. That this could be a quality measure for estimation is attributable to the vision model that was studied; the core model is an attractive mechanism for estimating a host of shape properties.

The perceptual descriptors of the physical quantities in the angiography experiment deserve highlighting. The figure-to-noise ratio, a measure proposed by Morse in his own dissertation research, appropriately specifies noise as an object-dependent property. Similarly, the effective blur scale also reflects object width in determining the overall amount of blur that is the combination of the image acquisition system's geometric blur and the visual system's object-proportional perceptual blur. These measurements were important as the stenosis estimates in the angiography experiment were made on different constricted vessel widths. This research on principle adopted these measures and expressed the former experimental design in those terms. Some evidence was adduced from the experimental results that suggested why that might have been a wise choice. Unfortunately, these recently-established principles have not yet been recognized by other researchers when characterizing noise and blur. Hopefully the work here can serve as a step toward proper specification of these quantities.

The human data that were collected, particularly from the angiography experiment, provide a valuable basis for subsequent improvements to the visual model. As work continues into developing and tuning the medialness or ridge components of the core model as well as the methods for making stenosis estimates that are derived from cores, it is this experimental data that can serve as the standard against which that progress is tested. When a model with a behavior that resembles these results is developed, an experimental investigation like the one developed in this dissertation could then be conducted again to test conclusively the proposed improvements to the model.

Finally, it is commendable in itself that this image quality measure is based on a model of the visual system. It was not the first time that principles of human vision have guided image analysis or evaluation. Nor were the results from this particular work especially predictive, though they were distinctly more so than previous tries. However, this research was arguably one of the most advanced applications to date of our knowledge of human vision in a medical imaging context.

Any research that is of an exploratory nature, and furthermore any study that marks the beginning of a research career, can always claim future directions that were identified during the process. Although this was mainly research into an image analysis method for computing quality, the human observer experiments used to establish the standard against which the model was tested could stand modification if they were to be redone. With belaboring the issues already pointed out in the appropriate places in the work, potential fixes for solidifying the human data include: an orthogonal design for the angiography experiment, estimates of intra-observer variability, common backgrounds in each experimental condition, and perhaps better methods for indicating the estimates.

The procedures that were proposed for producing an estimate of vessel stenosis or gap object width once a core for these objects was obtained represent the weakest theoretical component of what was called the model in this research. While the core model itself has physiological and psychophysical underpinnings, there is no way to know, once a perceptual representation of an object is constructed, how object characteristics are effectively extracted or combined from that representation. Short of further theoretical input, more experimentation could be done to systematically modify those procedures while simultaneously observing the results. It is an issue that will have to be addressed so long as the core or any other object representation scheme is utilized: for a quality measure, the object representation is only half the battle in modeling the performance of complex object-based tasks.

More research must be done to understand and characterize the sensitivity of the core model and the corresponding task estimates to the image characteristics that were manipulated in the experiments. The Monte Carlo approaches that were attempted in the angiography experiments allow extensive descriptive statistics of several aspects of the core representation in the presence of many iterations of noise or other manipulations such as blur, contrast, vessel and stenosis shape, and different backgrounds. These experiments proceeded with a best guess about medialness operators and methods for obtaining task information from cores. But there was little that was done in the way of testing or examining the model estimates prior to their comparison with the human data. More experiments on the core model alone must be carried out. It has always been said of the core that it is relatively insensitive to smaller scale degradations like blur and noise. Experiments could characterize, more fully than anything that was done here, under what conditions that is true of the model. Likewise, the experiments could tell more about the core representation as a function of the clinical variability in vessel shape properties. When the model is shown to have a predictably variable behavior, then it can be reexamined against what we now know about its human counterpart.

Detractors will complain that our knowledge of human vision is not sufficient to make this quality approach feasible. Rightfully, the experimental results from this work will fuel their fires. This process of using hypothesized principles of human vision in engineering applications that rely as heavily as this one did on the predictability of those principles may for some time meet with limited success. However, medical image interpretation by definition initiates with a visual experience. Inevitably, the list of known fundamental perceptual principles will grow. The production of pictures that humans view can only benefit from incorporating what we know about the visual process.