Visual search in medical images : a new methodology to quantify saliency

J. P. Rolland[1], K. Muller[2], and C. S. Helvig[1]

[1]Dept. of Computer Science, [2]Dept. of Biostatistics, University of North Carolina, Chapel Hill, NC 27516 (rolland@cs.unc.edu)

ABSTRACT

The quantification of object saliency in 2D greyscale images is of primary importance in medical imaging. Saliency, in this paper, refers to the detectability of objects of unknown location, an object being a feature of interest in an image. Object saliency is commonly measured by the speed of an observer at performing a detection task. The more salient an object is, the more quickly it can be detected. The questions raised in this paper are 1. whether the degree of saliency of an object can be solely predicted from its detectability as measured in a location known-exactly task, or whether factors such as geometry and context contribute to saliency in a more complex fashion; 2. whether the contribution of geometry and background complexity to saliency can be quantified. This paper focuses on the problem of detection of stenoses in simulated angiograms. Results from a first such study are presented. From those results, a new general methodology to measure saliency in 2D greyscale images was inferred and is presented.

1. INTRODUCTION

One of the challenges of assessing image quality for medical images stems from the complexity of the task to be performed and the complexity of the images involved. Medical imaging tasks, in the field of diagnosis, are generally of two types: classification and estimation.[1] A large body of literature on image quality for medical images has been focused on the detection of lesions of known location in uniform backgrounds,[2-5] and more recently, in lumpy backgrounds.[6-9] From such studies, researchers have investigated ways to predict detectability at unknown locations knowing observers' performance at known locations,[10-14] or investigated using eye-tracker technology how human observers search medical images.[15-16] In this paper, both classification (normal vs. abnormal), and estimation of the location of abnormalities, will be part of the task to be performed.

We chose angiography as the imaging modality due the rich and mathematically tractable geometry present in angiograms and the possibility to simulate remarkably realistic angiograms. Angiography is a radiological imaging modality that allows visualization of the vascular anatomy. High radiographic attenuation is achieved in the vessels of interest by the injection, prior to exposure with diagnostic radiation, of a radio dense contrast agent. One of the tasks performed with this modality is to detect or measure the degree and extent of bulges (aneurysms) and narrowings (stenoses) in blood vessels[17-19]. The presence of such abnormality may be indicative of a vascular disease; accurate diagnosis is essential in prevention of catastrophic vascular

0-8194-1784-X/95/$6.00

malfunction resulting from the rupture of an aneurysm or constricted blood flow in a major vessel.[20]

In this paper, we shall focus on the detection of stenoses in simulated angiograms. The simulation of an angiogram may be achieved by simulating individual blood vessels followed by the "artistic" summation of several of them. Clearly the real angiogram possesses many, branching blood vessels. We did not however find such a complicated simulation necessary for our studies, and the simplified approach that consists of adding individual blood vessels together leads to images that were actually mistaken to be real angiograms by several radiologists. Our general approach was to view three-dimensional (3D) blood vessels as general cylinders defined by a 3D medial curve and a corresponding width function associated with each point along the medial curve.[21] The simulation of an individual blood vessel consists then of the following steps 1) generating a 3D space curve that mathematically represents the 3D medial curve of the blood vessel, 2) defining a region of interest on the 2D projection of the 3D space curve to locate potential stenoses or aneurysms, 3) generating the 3D blood vessel from the 3D space curve and the parameters that characterize the blood vessel; e.g. width, type of defect if any, defect parameters such as the strength and length of the defect, 4) generating one or several 2D projections of the 3D blood vessel. A brief summary of the simulations is given in section 3.1, while more detail can be found in Rolland and Puff (1993).[22]

The research presented here has relevance not only to the medical community, but also to some of the vision community, especially researchers working in the field of shape perception. The task of detecting a narrowing in a blood vessel can indeed be thought of, more generally, as the detection of a change in the shape of an object. It is, for example, a question of current interest how and with what sensitivity do human observers perform shape discrimination. If the human visual system can be shown to respond differently to different geometrical features, independently of other features being present in the image, it is a significant finding towards better image understanding. Thus, the aim of this research is to also investigate some of those basic research questions.

This paper describes a quantitative investigation of shape discrimination in simulated angiograms. We report on an experimental visual search study and discuss a new methodology to quantify visual saliency in medical images.

2. RESEARCH QUESTIONS

The understanding of the problem of search in medical images is still today very much at its infancy. It is especially important to investigate due to the significant rate of misses reported in the clinic that are still not fully understood.[23]

Given the task of finding lesions in noisy backgrounds, some models of human performance have been derived to predict the detectability of lesions at unknown locations from the detectability of lesions at locations known-exactly.[14]

This research seeks to answer the following questions:

Can the degree of saliency of an object be solely predicted from its detectability measured in a location known-exactly task, or do factors such as geometry and context contribute to saliency in a more complex fashion?

Moreover, it is a challenge to characterize background complexity or context in non-simulated images, and to evaluate the impact of this complexity on the detection of abnormalities.[24]

Can the contribution of geometry and background complexity to saliency be quantified?

3. METHOD

Ultimately, only results on clinical images are relevant to the physicians, but because of the limited number of certified cases most likely available at an early stage of research and the time and difficulty involved in acquiring certified cases, realistic simulations may provide a successful springboard to clinical relevance. Depending on the type of structures imaged within the body and the nature of the patient (e.g. young vs. old; lean vs. fat), the images can range widely in complexity, with more or less complexity in the background (e.g. chest radiograph), texture (e.g. mammograms), and geometry (e.g. angiograms). This research focused on the problem of detection of stenoses in simulated angiograms since the complexity in such type of images is essentially one of geometry that can be mathematically characterized.

3.1 Angiogram simulations

Initially, a curve is generated which specifies the path of the vessel in three-dimensional space. The curve is specified by the Frenet formulas, mathematical specifications (or measurements) for the curvature and torsion of a 3D curve.[25] Given the curvature (κ) and torsion (τ) functions or values at each point along the arc length s of the space curve, the Frenet formulas can be used to compute the new Frenet frame field at a point $s+\delta s$. A curvature, κ, and torsion, τ, are chosen at each point along the space curve according to some specified statistical distribution about the previous values of κ and τ.[22]

Figure 1. *An example of (a) a simulated space curve and (b) the corresponding 2D projected blood vessel where an aneurysm was added.*

To simulate a 2D blood vessel, the 3D vessel volume whose position is determined by the space curve is first generated and then projected through the volume to form a 2D projection image. To generate the 3D blood vessel, a solid sphere of desired diameter is rolled along the space curve. A stenosis or aneurysm is generated along with the 3D blood vessel by varying the diameter of the sphere accordingly. Their shape is characterized by a Gaussian whose height contributes to specify the degree of aneurysm or stenosis and whose standard deviation specifies the length of the stenosis or aneurysm. The diameter of the rolling sphere is described as a constant plus a Gaussian function centered at the location of the stenosis or aneurysm. An example of a space curve and the resulting 2D blood vessel where an aneurysm was placed is shown in Figure 1 (a) and 1 (b), respectively.

3.2 Subjects

Nine subjects from 20 to 35 years old participated in the study, of which 5 were male and 4 were female. All subjects had 20/20 vision or corrected vision and no previous experience with reading angiograms or medical images.

3.3 Psychophysical study

This study aimed at investigating the effect of blood-vessel curvature at the location of the stenosis on detection performance in a visual search task. The search study involved one experiment. The experiment centered on 7 stimuli conditions summarized in Table 1. These conditions include 2 degrees of stenosis (45% and 95%) and three values of the mean 2D curvature along the stenosis in the 2D projection of the blood vessel for each degree of stenosis (0.0085, 0.0569, and 0.1817), computed from O'Neill (p75 ex14).[25] Note that the 2D curvature of a 2D space curve at a point along the curve is defined as the rate of change of the slope of the curve with arc-length at that point, and thus is unitless. These three values of curvature are referred in the table and hereafter as low, medium, and high curvatures, respectively. The length of the stenoses, when present, was constant and set to 20 pixels. The last condition listed in the first row of Table 1 corresponds to the absence of stenosis. This condition was used as a control condition to measure the number of false positives. Moreover, it helped mimic a more realistic clinical situation. The dependent measures were the reaction time to find the stenosis and the accuracy of localization of detected stenoses.

We generated 96 images per paired value of the two independent variables, in our case curvature (low, medium, high) and degree of stenosis (45%, 95%), and (96x3) images for the no-stenosis case. This yielded a total of 864 images for the experiment. A set of 36 images was used as a small training set prior to the experiment to get the observers familiar with the task. Each image was made of an average of 4 blood vessels in such a way that the mean ratio of filling of the images by the blood vessels was 16.5% while its standard deviation was 3.4%. Typical images are shown in Figure 2 (a) and (b), with a stenosis at a low point curvature and a high point curvature, respectively.

3.4 Task and Measures

Each image was presented for an infinite viewing time, to be terminated by the user. Subjects were asked to respond if they perceived a stenosis in the displayed image via a two button device. If "yes" was indicated, a cross appeared in the center of the image and subjects were asked to move the cross to the perceived location of the stenosis. Subjects were instructed to respond as quick as possible during the search task, while keeping the error rate below 5%. During the localization task no time limit was imposed on the subjects, and they were explicitly told that an accurate answer was important, while speed was irrelevant. Distinct auditory cues were provided after each trial, depending on the type of error: correct detection but false localization, or misses. In the case of incorrect localization, visual feedback was given to show where the stenosis was actually located, and where they had pointed. The 864 trials were performed in two sittings.

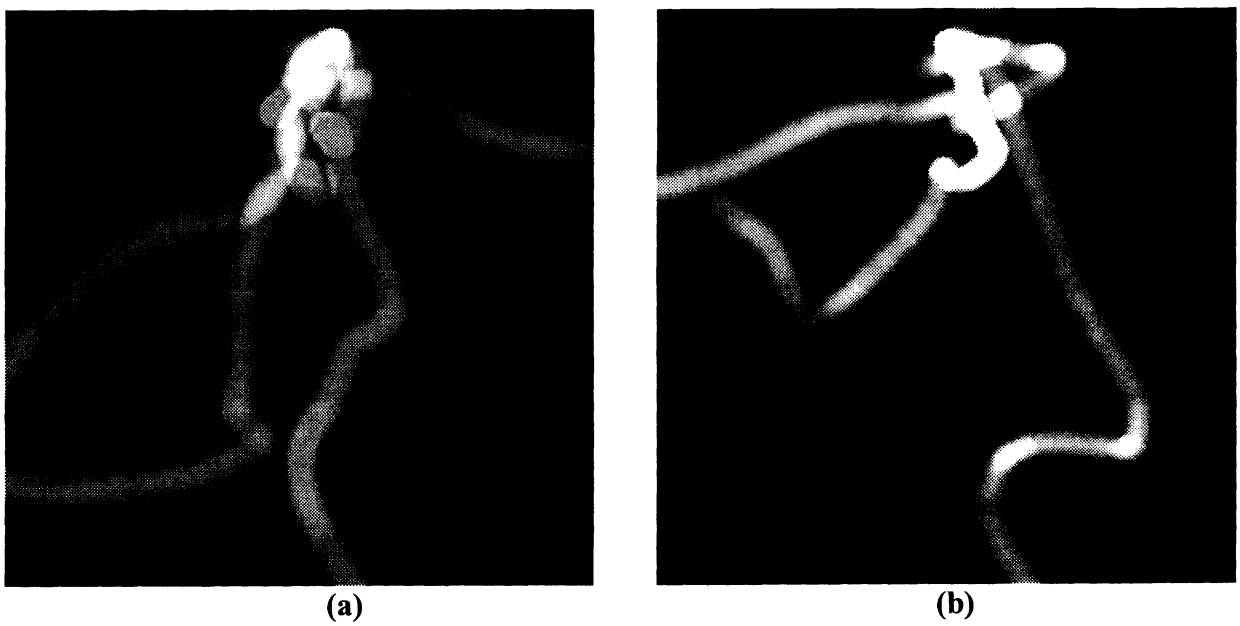

Figure 2. *An example of two stimuli used in the experiment: (a) an image with a stenosis at a low point of curvature; (b) an image with a stenosis at a high point of curvature.*

4. RESULTS

Median reaction time was computed for each of (96) or (96x3) trials in order to compute a robust estimate of performance. This approach is consistent with the results and recommendations of Ratcliff (1993).[26] The median reaction-time, averaged over the nine observers, is plotted as a function of the average 2D blood-vessel curvature along the stenosis for three degrees of stenosis: 0% or no stenosis, 45%, and 95% (Figure 3). The variations between observers are represented as ± 1 standard deviation. A repeated measures analysis of variance model was fitted to determine whether stenosis and curvature had an effect on median reaction time. The univariate approach to repeated measures was used, with the Geisser-Greenhouse corrected test. Increasing

stenosis tended to decrease reaction time, as expected. The differences between the median reaction-times at 45% and 95% stenosis for all three curvatures were all significant at the level p=0.0001. Results also show a significant interaction between stenosis and curvature (p=0.0498). The significant interaction indicates that the curve across curvature for 45% stenosis is not parallel to that for 95% stenosis. As can be seen in Figure 3, the interaction was ordinal. That means that, at each level of curvature, the effect of stenosis was in the same direction. Within each stenosis, the pairwise difference between median reaction-times was tested for (low-medium), (low-high), and (medium-high) curvatures. In these comparisons, the median reaction-times at high curvature are significantly different from both low and medium curvatures ($p \leq 0.0007$).

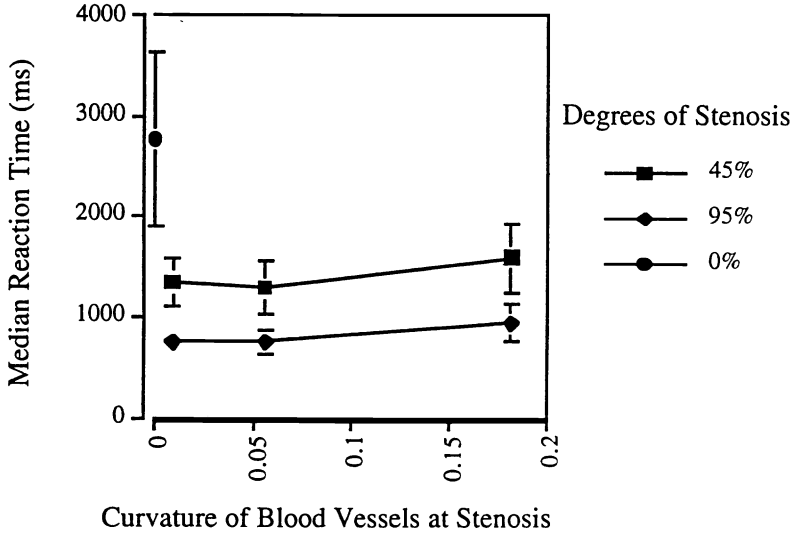

Figure 3. *Plot of the median reaction-times averaged over nine observers as a function of the 2D blood vessel curvature along the stenosis. Two degrees of stenosis are reported (45% and 95%) as well as the no-stenosis (0%) condition.*

Table 1: *Basic statistics on median reaction time by stenosis and curvature*

STENOSIS	CURV1	N	MEAN	STD	MAX	MIN
0	NONE	9	2762.0	859.54	3802.0	1342.5
45	LOW	9	1344.0	236.56	1704.5	1027.0
45	MED	9	1285.3	265.21	1782.0	908.0
45	HIGH	9	1583.7	345.20	2101.5	1253.0
95	LOW	9	754.9	116.55	916.5	563.0
95	MED	9	748.6	125.68	882.5	554.0
95	HIGH	9	927.7	184.75	1151.0	631.0

We also analyzed the data in terms of the number of misses and false alarms per degree of stenosis and curvature. A trial is counted as a miss when the image contained a stenosis but the observer either responded that no stenosis was present, or responded that a stenosis was present but located the stenosis outside of a circle of 20 pixels around the center of the stenosis. A value of 20 pixels was chosen to correspond to the extend of the simulated stenoses themselves. For the 45% stenoses, the total number of misses, averaged over the nine observers, was 7.8, 6.7, and 17.3, for low, medium and high curvature, respectively. Those numbers correspond to percent misses of 8%, 6.9% and 18% by dividing the averaged numbers of misses by 96 which is the total number of images presented per category. For the 95% stenoses, the averaged number of misses was 0.3, 0.3, and 1.1, yielding percent misses of 0.3%, 0.3%, and 1%, for low, medium, and high curvatures, respectively. The total number of false alarm averaged over the nine observers was 6.8, which corresponds to a percent false alarms of 2.4%.

5. DISCUSSION: A NEW METHODOLOGY TO QUANTIFY SALIENCY

The results of the experiment described above seem to indicate that low-curvature points along the blood vessels are more salient than high-curvature points. However, because of the significant increase in the median reaction-time for the high-curvature points only, while low and medium curvatures yielded non-significant differences, one could postulate that, even with such a high degree of stenosis as 95%, a narrowing of the blood vessel at a high-curvature point is perhaps less detectable to start with than a narrowing at a low or medium curvature point. We mean by less detectable here that, even if the potential location of the stenosis was indicated, the observer would rate his certainty to see the stenosis at a lower level for a high-curvature point than for a low- or medium- curvature point.

Those results spawned a new approach to measuring feature saliency in medical images. This approach was independently arrived at by Palmer (1994).[27] It consists in measuring and comparing saliency of features which have been set equally detectable for a location known-exactly task. We thus propose a two step procedure to measure feature saliency. The first step consists of measuring the detection thresholds of those features, in our case a set of differently shaped stenoses, in a location known-exactly task. The second step consists in measuring the reaction time of finding one of those features in a visual search task.

A pilot study aimed at finding stenosis amplitude-thresholds was first conducted with two observers using a yes/no paradigm with 50% catch trials. Under this investigation, this non-criteria free method did not seem to be able to give us a clean measure of the thresholds needed without making several assumptions which would have required extensive experiments to test. Such a method, in any case, would give us at best the equivalent of one point on the ROC curve. A standard ROC study using 8 observers is currently under way to measure stenosis-amplitude thresholds, with early data looking very promising.

Using the measured stenoses amplitude thresholds for differently shaped stenoses, we propose in a future study to repeat the visual search task for stenosis amplitudes that are equal to the detection thresholds and its multiples up to 100% stenosis.

We speculate that the effect of background complexity on both detectability and saliency can also be accounted for by extending the location known-exactly ROC study described above for blood vessels embedded in their own context. Much work remains to be done before further conclusions or insights can be drawn.

6. CONCLUSION

We present results on a visual search investigation that aims at studying the ability of the human observer to detect a change in shape of a 2D greyscale object embedded in a context of similar objects. Human observers were asked to detect stenoses in simulated blood vessels embedded in simulated angiograms. Results from a visual search experiment were presented and indicate that stenoses at high-curvature points along the blood vessels are less salient than stenoses at low-curvature points. More importantly, those results also spawned a useful approach to measuring feature saliency in medical images.

7. ACKNOWLEDGMENTS

We are grateful to Christina Burbeck and Stephen Pizer for their stimulating discussions about this research. We also thank John Roades for his suggestions concerning the implementation of the rolling sphere algorithm used in the angiogram simulations and Robert Gardner for sharing his expertise and insight in differential geometry. Many thanks go to Haesook Kim and Deborah Glueck for their help with the data analysis. Finally we thank our human subjects for their kindness and patience in participating. The research reported here was supported by NIH grant #PO1 CA47982.

8. REFERENCES

1. H.H. Barrett, "Objective assessment of image quality: effects of quantum noise and object variability," J. Opt. Soc. Am. A, 7 (7), pp. 1990.

2. Whalen, A.D. Detection of signals in noise (Academic, New York), 1971.

3. A.E. Burgess, R.F. Wagner, R.J. Jennings, and H.B. Barlow, "Efficiency of human visual signal discrimination," Science 214, pp. 93-94, 1981.

4. P.F. Judy, and R.G. Swensson, "Lesion detection and signal-to-noise ratio in CT images," Med. Phys., 8, pp. 13-23, 1981.

5. R.F. Wagner, D.G. Brown, "Unified SNR analysis of medical imaging systems", Phys. Med. Bio., 30 (6), pp. 489-518, 1985.

6. K.J. Myers, J.P. Rolland, H.H. Barrett, and R.F. Wagner, "Aperture optimization for emission imaging: effect of a spatially varying background," J. Opt. Soc. Am. A. 7, pp. 1279-1293, 1990.

7. J.P. Rolland and H.H. Barrett, "Effect of random background inhomogeneity on observer detection performance," J. Opt. Soc. Am. A. 9, pp. 649-658, 1992.

8. H.H. Barrett, J. Yao, J.P. Rolland, and K.J. Myers, "Model observers for assessment of image quality," Proc. Natl. Acad. Sci., 90, pp. 9758-9765, 1993.

9. A.E. Burgess, "Statistically defined backgrounds: performance of a modified nonprewhitening observer model," J. Opt. Soc. Am. A., 11(4), pp. 1237-1242, 1994.

10. L. W. Nolte, and D. Jaarsma, "More on the detection of one of M orthogonal signals," Journal of the Acoustical Society of America, 41, pp. 497-505, 1967.

11. S.J. Starr, C.E. Metz, L.B. Lusted, and D.J. Goodenough, "Visual detection and localization of radiographic images," Radiology, 116, pp. 553-538, 1975.

12. A.E. Burgess, and H. Ghandeharian, "Visual signal detection. II. signal-location identification," J. Opt. Soc. Am. A., 1, pp. 906-910, 1984.

13. Pelli, D.G., "Uncertainty explains many aspects of visual contract detection and discrimination," J. Opt. Soc. Am. A. 2, pp. 1508-1532, 1985.

14. R.G. Swensson, "Measuring detection and localization performance," Information Processing in Medical Imaging, H.H. Barrett and A.F. Gmitro, Springer-Verlag, pp. 525-541, 1993.

15. H.L. Kundel, C.F. Nodine, and E.A. Krupinski, "Searching for lung nodules: visual dwell indicates locations of false positive and false negative decisions," Investigative Radiology, 24, pp. 472-478, 1989.

16. C.F. Nodine, H.L. Kundel, L.C Toto, and E.A. Krupinski, "Recording and analysing eye-position data using a micro-computer workstation," Behavior Research Methods, Instruments, and Computers, 24, pp. 475-485, 1992.

17. T. Sandor, and R.G. Swensson, "Evaluation of observer performance in detecting blood vessels on simulated angiographic images," Med. Phys. 5(5), pp. 380-386, 1978.

18. J. Tobis, O. Nalcioglu, L. Iseeri, et al., "Detection and quantitation of coronary artery stenoses from digital substraction angiograms compared with 35 millimeter film cineangiograms," Am. J. Cardiol 54, pp. 489-496, 1984.

19. P. Jaques, F. DiBianca, S. Pizer, F. Kohout, L. Lifshitz, and D. Delany, "Quantitative digital fluorography: computer vs. human estimation of vascular stenoses," Investigative Radiology, 20, pp. 45-52, 1985.

20. J. Ambrose, S.L. Winters, A. Stern, A. Eng, L.E. Teichholz, R. Gorlin, and V. Fuster, "Angiographic morphology and the pathogenesis of unstable angina pectoris," J. Am. Coll. Cardiol. 5(3), pp. 609-616, 1985.

21. H. Blum, and R.N. Nagel, "shape description using weighted symmetric axis features," Proc. of the IEEE Computer Society Conference on Pattern Recognition and Image Processing, pp. 167-180, 1977.

22. J.P. Rolland, and D.T. Puff, "Angiogram simulation software documentation," University of North Carolina Department of Computer Science, Technical Report TR93-018, 1993.

23. R.G. Swensson, S.J. Hessel, and P.G. Herman, "Omissions in radiology: faulty search or stringent reporting criteria," Radiology, 123 (3), pp. 563-567, 1977.

24. G. Revesz, H.L. Kundel, and M.A. Graber, "The influence of structured noise on detection of radiologic abnormalities," Invest. Radiol. 9, pp. 479-486, 1974.

25. O'Neill, Barrett. Elementary Differential Geometry. Orlando: Academic Press, 1966.

26. R. Ratcliff, "Methods for dealing with reaction time outliers," Psychophysical Bulletin, 114 (3), pp. 510-532, 1993.

27. J. Palmer, "Visual search latency: the influence of target distractor discriminability on the magnitude of set-size effects," Investigative Ophthalmology and Visual Science, 35(4), (abstract), 1994.