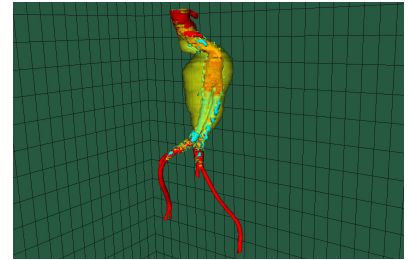


# Semi-automated Measurement of Anatomical Structures Using Statistical and Morphological Priors

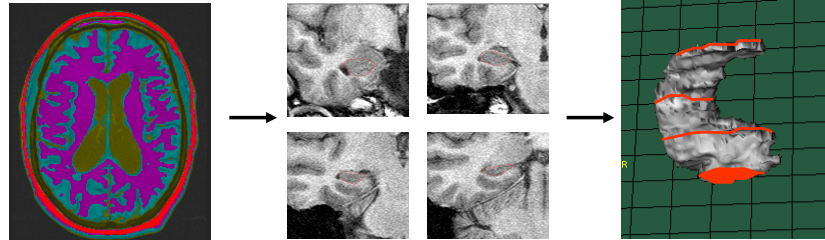
Edward Ashton and Tong Du  
VirtualScopics, LLC, Rochester, NY



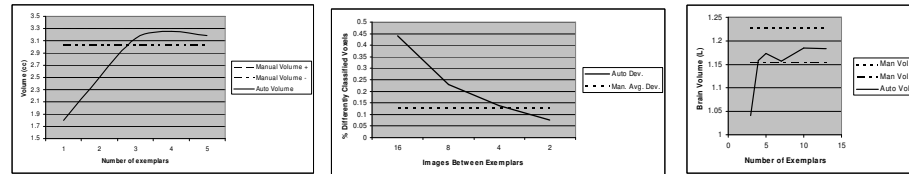
**BACKGROUND:** Accurate identification and measurement of various anatomical structures is a vital tool both for surgical planning and for evaluation of disease progression and patient response to therapy for numerous diseases. Measurement of hippocampal volume is an important endpoint for diagnosing and monitoring both intractable temporal lobe epilepsy and Alzheimer's disease. Identification of the aorta and associated vessels and measurement of various related parameters are vital tools for evaluation of abdominal aortic aneurism progression and response to treatment. Current standard methods for obtaining these data points are largely manual and subjective, and are therefore both error-prone and subject to inter- and intra-operator variability. In addition, manual tracing of structures such as the vascular system, which may appear on up to 800 images in a single study, requires both considerable expertise and a great deal of time.

**MEASUREMENT TECHNIQUES:** The anatomical feature delineation technique evaluated in this work makes use of three types of information: (1) A statistical description of the various tissue types present in the images. This information is obtained automatically through the use of a maximum likelihood classifier. (2) A statistical description of the tissue or tissues of interest. This information is obtained by making use of an anatomical atlas or user input – typically one or more seed regions or exemplars. The exemplars will ideally take the form of an identification of the structure of interest on one or more non-adjacent images. (3) A morphological description of the structure of interest. This information can be taken from an atlas-derived *a priori* shape model in the case of well-defined and localized structures such as the hippocampus. If such information is not available, a three-dimensional surface is fit to the provided exemplar regions. In the case of the abdominal vascular system, a surface is fit to the outer boundary of the lumen, and another is fit to the outer vessel wall. Other tissue types, such as thrombus and calcifications, are assumed to exist in the region between these two boundaries. These data serve as the inputs to a second-pass maximum likelihood classifier, which identifies the boundaries of the structure or structures of interest. This is illustrated in Fig. 1. Information flow for the full system is shown in Fig. 2.

**ABSTRACT:** Rapid, accurate and reproducible delineation and measurement of arbitrary anatomical structures in medical images is a widely held goal, with important applications in both clinical diagnostics and, perhaps more significantly, pharmaceutical trial evaluation. This process requires the ability first to localize a structure within the body, and then to find a best approximation of the structure's boundaries within a given scan. Structures that are tortuous and small in cross section, such as the hippocampus in the brain or the abdominal aorta, present a particular challenge. Their apparent shape and position can change significantly from slice to slice, and accurate prior shape models for such structures are often difficult to form. In this work, we have developed a system that makes use of both a user-defined shape model and a statistical maximum likelihood classifier to identify and measure structures of this sort in MRI and CT images. Experiments show that this system can reduce analysis time by 75% or more with respect to manual tracing with no loss of precision or accuracy.



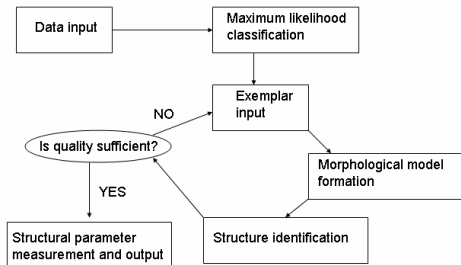
**Figure 1:** Preliminary processing steps: Automated statistical characterization, exemplar generation, and 3D surface fitting.



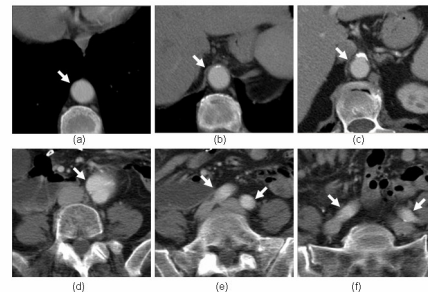
**Figure 3:** Accuracy results for hippocampus measurement, vessel lumen identification, and brain volume measurement. For the hippocampus and brain measurement the statistic of interest is deviation from manually measured volume. For the lumen identification the statistic of interest is differently classified voxels. In all three cases, results within one standard deviation of manual are achieved with approximately four images between exemplars, yielding a time savings of 80% with no loss of accuracy.

	Repeat 1	Repeat 2	Repeat 3	Repeat 4	Mean	Std. Dev.	Coef. Var.
Subject 1	3.143	3.284	3.183	3.268	3.22	0.06	2.1%
Subject 2	3.34	3.351	3.379	3.154	3.306	0.10	3.1%
Subject 3	3.219	3.21	3.27	3.259	3.24	0.03	0.9%
Subject 4	2.647	2.836	2.79	2.748	2.755	0.08	2.9%
Subject 5	3.069	3.179	3.132	3.179	3.14	0.05	1.7%

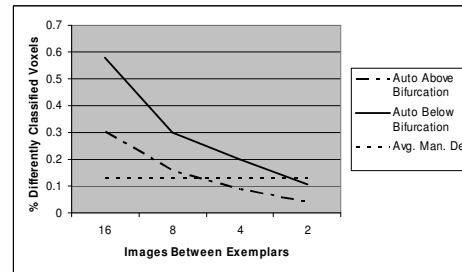
**Table 1:** Reproducibility results for the hippocampus experiment. Mean coefficient of variability is 2.1%.



**Figure 2:** Information flow for the analysis system. If initial quality is insufficient, additional exemplars can be defined to further refine the morphological model.



**Figure 4:** Images from an abdominal CT scan. Separation between (a) and (c) is 4.8cm. The aorta changes very little in this region, allowing a close fit with only 2 exemplars. Separation between (d) and (f) is also 4.8cm. In this region the vasculature is smaller and more tortuous, requiring a greater number of exemplars for an accurate surface fit.



**Figure 5:** Classifier performance on lumen, above and below the aortic bifurcation. As expected, performance above the bifurcation (Fig. 4 (a) – (c)) is significantly better than that below the bifurcation (Fig. 4 (d) – (f)).

**EXPERIMENTAL PROCEDURE:** Three experimental applications of this technique were tested in this work. In the first, the process was used to identify, delineate and measure the hippocampus in T1 weighted MRI images of normal volunteers. In the second, the process was used to identify, delineate and measure (separately) the lumen and surrounding thrombus in CT images of patients suffering from abdominal aortic aneurisms. In the third, the process was used to measure the total brain volume in a multiple pulse sequence MRI head scan. All experiments sought to quantify three variables describing system performance: processing time, accuracy with respect to manual tracing, and reproducibility in terms of coefficient of variability over repeated measurement of the same anatomy.

The first two variables are intertwined, since the goal of the system is to achieve optimal processing time while producing results statistically indistinguishable from those obtained by manual tracing. Preliminary experiments were carried out in order to define a standard deviation for manual measurement of the structures under consideration. Automated results were considered to be indistinguishable from manual results if they fell within one standard deviation (as determined by these experiments) of the corresponding manual measurements.

**RESULTS:** In order to establish a gold standard and an associated error margin for each experiment, the structures of interest in each subject were identified by four expert analysts using a computer-aided manual tracing process. The experiment was intended to determine: (1) How many exemplars were required to produce an automated measurement that was statistically indistinguishable from a manual one? (2) What was the time savings associated with this process, as compared to manual tracing? (3) What was the reproducibility of the automated process?

In order to answer the first question, the structures of interest were measured, using a varying number of exemplars for morphological model formation. These results were compared to manual measurements of the same structure. The results of this experiment are given in Figure 3.

In the second phase of the hippocampus experiment, the right hippocampus of each of five subjects was analyzed four separate times. The intent in this case was to establish the reproducibility of this technique. Results of this experiment are given in Table 1. Clearly the reproducibility of this system is significantly better than that of manual tracing.

**CONCLUSIONS:** The analysis system developed during the course of this work appears to have broad applicability, working more or less equally well on single pulse sequence MR, multiple pulse sequence MR, and CT, and on structures ranging from roughly ovoid to long and tortuous. System accuracy is dependent on the accuracy of the morphological model, whether it is derived from an anatomical atlas or from exemplar regions.

It should also be noted that all results shown in Fig. 3 were derived from evenly spaced exemplars. It can be demonstrated that intelligent selection of exemplars, with more placed in tortuous regions and fewer in relatively regular regions, can produce the same results with significantly fewer exemplars. This is illustrated in Figs. 4 and 5.