Roberto Macchiarelli & Gerd-Christian Weniger (eds.):
Pleistocene Databases: Acquisition, Storing, Sharing.
Wissenschaftliche Schriften des Neanderthal Museums 4, Mettmann 2011.

THE OPEN RESEARCH SCAN ARCHIVE (ORSA): A MASSIVE OPEN-ACCESS ARCHIVE OF RESEARCH QUALITY COMPUTED TOMOGRAPHY (CT) SCANS

Janet Monge and P. Thomas Schoenemann

Abstract

The Open Research Scan Archive (ORSA) is a large and expanding archive of high resolution, research quality computed tomography (CT) scans of osteological specimens, covering both human and non-human crania and postcrania. At present, the archive contains over 4000 scans. The archive is open to any researcher presenting a clear research proposal. A searchable online database containing details of the available specimens is available, and requested scans can be delivered via a specialized web interface or DVD disk. ORSA is also unique in attempting to encourage open access of scan collections generally, by requiring researchers obtaining scans to commit to contributing their own scans back to the general research community (immediately if the researcher has scans available, otherwise in the future). This can be done through ORSA or via some other mechanism.

Keywords: Computed Tomography, CT Scans, Skeletal Material, Online Database, Web Access, Data Sharing

Introduction

Research in anthropology, skeletal biology, and paleontology has been greatly enhanced by advances in 3-dimensional (3D) imaging techniques. Once specimens have been CT scanned, a large number of detailed anatomical studies can be accomplished without having to handle the actual specimens again. CT scans can now be easily sent via internet all over the world. A large and expanding set of useful software tools, many of them freely available as open source, allow complex mathematical analyses of shape and form to be applied to these scans, both at the individual and population levels. This has greatly enhanced both the types and sophistication of research projects that are possible. Recognition of this has lead to the widespread scanning of fossil specimens (though access to these scans is unfortunately often limited at present). Recent fossil finds in paleoanthropology are invariably CT scanned.

In light of this, the ORSA project was conceived in 2001 with the idea of creating a large CT archive that would be open to all researchers. The idea was to make available a large comparative dataset of skeletal specimens which could allow for greater standardization of comparisons across research projects. This was possible in part because of the unique skeletal collection available at the University of Pennsylvania Museum of Archaeology and Anthropology, one highlight of which is the Morton collection of human crania. This is a collection of over 1800 specimens from all over the world, and is probably the broadest such collection in existence. Scanning began in 2002 and has continued to the present. The archive now contains over 4000 scans, and is still growing. The majority of the funding for this project has come from the U.S. National Science Foundation, with seed money and infrastructure help from the University of Pennsylvania. More recently, Indiana University has provided invaluable information technology infrastructure support as well.

The database is designed to continually grow. As each new scan is obtained it is made available online in order to maximize its usefulness to researchers worldwide. To date, well over 100 researchers and scholars have used these CT scans as part of their own research agendas.

Research use of CT

The resolution possible from CT scanners is significantly finer than the degree of accuracy obtainable from direct measurements of the original specimen using calipers. Combined with currently available 3D visualization software, it is possible to take

any measurement from the CT scans themselves. For most types of research this eliminates the need to touch the actual specimens again minimizing the likelihood of damage. It also means that, given the appropriate software, research on these specimens can be done anywhere. CT scans reconstruct internal structures that are not studied externally. Studies of the evolution of bipedalism, for example, have included analyses of the orientation of the inner ear. In the absence of CT, this would not be possible without destroying specimens.

Recent advances in 3D image analysis for clinical research allow for new and more powerful studies of complex geometry of the skull. For example, algorithms have been developed that allow several brains to be morphed into a common coordinate system for use in functional brain imaging. This allows more accurate comparison of brain activation across individuals with brains of different sizes and shapes. We have begun applying these algorithms to create high-resolution 3D maps of individual variability in morphology. This information is routinely discarded in functional imaging studies once individual functional scans are morphed into the same brain space; however, it is an extremely rich source of information that can be used to study shape in much more sophisticated ways than has been possible previously.

GOALS AND ASPIRATIONS

The express purpose of the collection is to facilitate research in biology, skeletal biology, anthropology, medicine, and other related disciplines. It is hoped that the database will become a clearinghouse for research scan data of all kinds, including CTs of fossil specimens. This database will allow us to better interpret such fossils by allowing us to place them in comparative perspective with other specimens. The following Institutions have already contributed specimens to the collection: University of Pennsylvania Museum of Archaeology and Anthropology, the American Museum of Natural History, the Smithsonian Institution, the Columbia University Department of Anthropology, the Museo Nacional de Ciencias Naturales, CSIC, the Hull York Medical School, the Florida Atlantic University, the University of Michigan - Ann Arbor, the Academy of Natural Sciences, the Mutter Museum, College of Physicians of Philadelphia.

ORGANIZATION OF ORSA - SPECIMENS IN THE DATABASE

The Samuel George Morton collection is housed and curated at the University of Pennsylvania Museum of Archaeology and Anthropology. The original Morton collection is composed of approximately 1200 human crania, most without mandibles, and collected from both archaeological and recent contexts (1820's to 1851). After Morton's death, his student, J. Aitken Meigs, continued with the collection which totals approximately 1800 crania. In addition, the database includes scans of three modern human crania along with scans of their matching plaster endocasts created by Ralph Holloway at Columbia University.

Several orangutan (*Pongo pygmaeus*) crania in the database come from the Harrison and Hiller University of Pennsylvania Museum of Archaeology and Anthropology expedition to Borneo late in the 19th Century. All specimens were wildshot, prepared in Borneo, and shipped to Philadelphia. Most of the chimpanzee (*Pan troglodytes*) specimens are from the American Museum of Natural History, from the von Lushen collection. The database also includes non-human primate specimens from the University of Pennsylvania Museum of Archaeology and Anthropology, which were originally obtained as part of a gift from the Academy of Natural Sciences, Philadelphia. All specimens date from the 1880s-1890s.

ORGANIZATION OF ORSA - ACCESS AND FACILITIES

The ORSA CT image archive is stored on computers both at the University of Pennsylvania Museum of Archaeology and Anthropology and at Indiana University.

An online informational database allows researchers to search the collections

based on user-defined criteria (see below). The website address is: http://plum.museum.upenn.edu/~orsa/Welcome.html. Researchers either come to work on the data at our labs, obtain scans via a separate web access portal that holds the actual scans, or have their scans sent to them on CD or DVD.

Organization of ORSA - Access policies

The ORSA project is unique among CT archives in its encouragement of data sharing. This is partly due to the fact that it takes a museum rather than a research perspective on specimen access. Museums typically take as their fundamental mandate the idea that their collections are in some important sense owed to (if not literally owned by) the public, and that their mission is specifically to provide access to as wide a range of interested people as possible. Researchers have typically taken the opposite perspective: that because they collected the data - often under difficult field circumstances - they should therefore have the right to control access until they have completed the most important and interesting research on them. Given that ORSA's collection is based on museum specimens, and given that a public funding institution (the National Science Foundation) provided the resources for ORSA (though admittedly this is also true of many research projects), the idea of open and free access is taken very seriously by ORSA. The only requirement for obtaining access is that the researchers provide a clear research proposal.

ORSA also attempts to encourage data sharing. This is done by requiring - as a condition of obtaining access to ORSA scans - that potential users agree to donate scans back to the archive for others to use. Any access restrictions placed on these donated scans are respected (typically: requiring interested researchers to obtain permission from the original donator). The web portal used by ORSA provides a convenient mechanism for uploading scans to the database. ORSA realizes that not everyone with a great research idea has scans to donate. In particular, graduate students at the beginning of their careers often have nothing to donate (let alone to do research with in the first place). In such cases individuals are instead required simply to pledge they will donate future scans they might obtain back to ORSA. The hope is that their research with ORSA specimens will allow them to obtain future funding for more scanning, or open doors for them in research institutions that might have scans that could be donated.

The specific ORSA policies regarding access of scans are as follows:

1. Any material obtained from this archive (scans, data, reports, images, etc.) may be used for academic, research, and educational purposes ONLY. Use of any of this material for commercial purposes is strictly prohibited.
2. Any discussions, presentations, or publications of any work done with material obtained from this archive must include acknowledgment of the University of Pennsylvania Museum of Archaeology and Anthropology, the Open Research Scan Archive at Penn, as well as Janet Monge and P. Thomas Schoenemann. We also have some scans from specimens from other institutions. These institutions would need to be acknowledged in addition, if their specimens are used.
3. Citation details of any discussions, presentations, or publications of any work done with the CT's must be sent to us as well, so that we have a detailed record of how extensively the archive is being used (thereby helping us secure possible future funding to help expand the collection for the benefit of everyone).
4. Data and scans obtained from this archive must not be forwarded to anyone else. Instead, other researchers interested in doing research on the CT files should contact us directly, and we will be happy to accommodate them. Again, this helps us keep track of how extensively the data is being used. Researchers who donate scans to the archive of course retain the right to independently send their scans to anyone they choose (without contacting ORSA).
5. In the spirit of open access to data, we require that anyone with image datasets of their own contribute these to the Open Research Scan Archive in exchange for ob-

taining scans currently in our archive. Any restrictions you wish to place on these donated images will of course be respected.
6. Requested CT files will be burned to a CD or DVD. A small fee will be charged to cover the time and postage.
7. We welcome visiting researchers to use the image processing tools available in our lab. (Interested individuals are encouraged to let us know if this is of interest, and we then arrange the details).
8. We can also arrange to have specimens scanned at Penn, if this is of interest. Our current cost is approximately $10 per scan, depending on time and availability of the scanner. Depending on the level of interest in the specimen(s), this fee may be waived (please contact us for more details if you have a specimen you need scanned).

ORGANIZATION OF ORSA - DATA STORAGE

The ORSA archive actually consists of two separate databases. As mentioned above, one of these databases contains information about the specimens themselves, including museum records and any information that is known about the specimen (e.g., species, age, sex, provenience, etc.). This database is currently maintained using Filemaker, which provides a simple webserver interface for this information, allowing interested researchers to search for possible specimens of interest. This information database is accessed through the following website: http://plum.museum.upenn.edu/~orsa/Welcome.html.

The scans themselves are currently stored in DICOM format, in a picture archiving and communication system (PACS), using an Apple OS X open source clinical image and object management system called OsiriX. This runs on a workstation with 2TB storage, which is backed up hourly to an external 4TB hard disk. Additional backup is provided by forwarding new studies automatically to a mirror archive located at Indiana University. OsiriX stores the images, allows for easy searching for specific scans, and creates a web interface simplifying the distribution of scans to researchers, as well as simplifying the creation of archive mirrors (copies of the archive at other institutions). Web download is provided through the Indiana University mirror, because this server is placed close to the main internet portal to Indiana University, which allows for very fast download speeds. The hope is to expand mirrors to other interested institutions, thereby further increasing access. Web access to specific requested scans is easily set up after researchers provide a description of their research plans, a list of specimens they would like to work on, and agree to the access policies (see above).

RESEARCH USE OF ORSA TO DATE - STUDIES BASED ON ORSA DATA

Several studies made possible by scans from the ORSA database have been reported to date:

- The relative accuracy of handmade plaster replicas of the endocranium (inside of the braincase) with respect to the actual endocranial surface, as determined from CT scans of skulls. The results show that handmade replicas are accurate to within about 2 mm either way of the CT derived surface, though the error that does exist tends to have a very idiosyncratic pattern from specimen to specimen (Avants et al. 2004; Schoenemann et al. 2007a).
- An assessment of how ape endocrania differ from human endocrania is being studied by morphing ape and human images together. The differences between individual specimens is then characterized in 3 dimensions, on a voxel-by-voxel basis, maximizing the chance of extracting meaningful information from fossil endocasts (Schoenemann et al. 2011).
- A variety of methodological studies have been done, including the objective mathematical assessment of surface curvature for skeletal (specifically endocranial) surfaces (Avants et al. 2005), studies of different methods for estimating cra-

nial capacity (Lewis et al. 2004), studies of sex and age differences in populations of specimens (Schoenemann et al. 2007b), studies of asymmetries in endocranial shape in apes and humans (Schoenemann et al. 2008), and the creation of a statistical atlas of modern human cranial derived from non-rigid morphing techniques (Schoenemann et al. 2009).

- A geometric morphometric study of the middle cranial fossa in modern and fossil humans by Bastir et al. (2008), using scans from ORSA, confirmed that the middle cranial fossa of *Homo sapiens* is significantly different from *H. neanderthalensis*, *H. heidelbergensis*, and *Pan troglodytes*.
- CT analysis can be used for the analysis of trauma on archaeological specimens. In the 1920s, two skulls from the Royal Tombs or Ur were excavated in situ due to their fragile nature. The skulls were CT scanned and reconstructed to show differences in the pattern of perimortem and postmortem bone fracture (Baadsgaard et al. in press).
- The Penn Museum curates approximately 50 mummies. In conjunction with the American Association of Anatomists and the Anatomical Record, results of the CT scans of the Egyptian mummy are planned to be presented at a symposium at the Penn Museum – "The Anatomy of a Mummy" – February 26, 2012. Along with the Editor-in-Chief of the Anatomical Record, a full issue of the journal will be devoted to the CT analysis of mummies from all over the world.

Research use of ORSA to date - Ongoing research using ORSA scans

Several studies are currently being addressed with ORSA scans. The following are highlights of this work:

- Comparative assessment of the placement on the cranial base of the foramen magnum (where the spinal cord exits the base of the skull) in bipedal humans vs. non-bipedal apes is being assessed through the use of ORSA scans.
- Explorations of the functional purpose of large supraorbital tori (brow ridges) is being assessed by exploring associations between variation in size of supraorbital tori and other parts of the skull, such as the masticatory system. This study is using morphing techniques on a large sample of modern human crania, leading to a description of variability on voxel-by-voxel basis. One can then map the extent to which variability at each point correlates with other landmarks or morphological features.
- Research into the possible functions of the maxillary sinus is being pursued by Lauren Butaric at Texas A&M University using scans from ORSA. She is specifically interested in whether exists for some specific physiological function, or is just an architectural byproduct.
- Researchers at the Centro Nacional Patagonico (part of CONICET), led by Rolando Gonzalez-José are studying craniofacial variation and evolution in New World human settlement, hunter-gatherer populations. In exchange for access to specimens from ORSA, these researchers have agreed to contribute 30 their own CT scans of modern Brazilian individuals to ORSA for future use by others.

Publications advertising ORSA

Information and announcements about ORSA have been published in the American Journal of Physical Anthropology (Monge et al. 2004), and the British Institute of Radiology Newsletter (Schoenemann et al. 2008b). Posters have also been presented at the Meetings of the American Association of Physical Anthropologists (April 2004), as well as the Annual Meeting of the Natural Science Collections Alliance (June 2003). In addition, numerous media reports have been highlighted various projects made possible by ORSA also: WHYY Radio, December 15 2008, "Scanning skulls for their secrets," and a Penn Museum press release (picked up by the Associated Press and various other local media): "Two 4600 Year Old Skulls from Famous Excavations at Royal Tombs of Ur,

Iraq Traveled to the Hospital of the University of Pennsylvania For CAT Scans Sunday morning, April 15", April 2007.

ACKNOWLEDGEMENTS

This project is funded by the National Science Foundation (USA) (Grant number: 0447271). We also thank the University of Pennsylvania Research Fund and the Institute for Research in Cognitive Science for their generous support; and Dr. Nick Bryan and Felicia Jefferson for allowing us to use the CT facility at the Hospital of the University of Pennsylvania. In addition, the University Information Technology Service (UITS) at Indiana University, in particular Kurt Seiffert, manager UITS Research Storage, have provided critical technical and hardware support, including providing the server and backup storage facilities for the main ORSA web portal. We would also like to thank Dr. Ian Tattersall (American Museum of Natural History), Dr. Kenneth Mowbray, and Dr. Ralph Holloway from Columbia University, for allowing specimens in their care to be scanned, and to all the researchers who have contributed scans to the collection to date.

Janet Monge
Adjunct Associate Professor
Department of Anthropology
University of Pennsylvania
USA
jmonge@sas.upenn.edu
and
Associate Curator
Physical Anthropology
University of Pennsylvania Museum
USA

P. Thomas Schoenemann
Associate Professor of Anthropology
Department of Anthropology
Indiana University
USA
toms@indiana.edu

References cited

Avants, B., Gee, J., Schoenemann, P. T., Monge, J., Lewis, J. E. & Holloway, R. L. (2004). Validation of plaster endocast morphology through 3D CT image analysis. *American Journal of Physical Anthropology* 123 (Supplement 38): 56.

Avants, B. B., Gee, J. C., Schoenemann, P. T., Monge, J., Lewis, J. E. & Holloway, R. L. (2005). A new method for assessing endocast morphology: calculating local curvature from 3D CT images. *American Journal of Physical Anthropology* 126 (Supplement 40): 67.

Baadsgaard, A., Monge, J., Zettler, R. & Cox, S. (in press). Human sacrifice and intentional corpse preservation at Ur. *Antiquity* 85 (327): 27-42.

Bastir, M., Rosas, A., Lieberman, D. E. & O'Higgins, P. (2008). Middle Cranial Fossa Anatomy and the Origin of Modern Humans. *The Anatomical Record* 291: 130-140.

Lewis, J. E., Schoenemann, P. T. & Monge, J. (2004). Endocranial capacity estimated from 3-D CT: Methodological issues. *American Journal of Physical Anthropology* 123 (Supplement 38): 135.

Monge, J., Schoenemann, P. T., Lewis, J. & Glotzer, D. (2004). The CT Database at the University of Pennsylvania Museum. *American Journal of Physical Anthropology* 123 (Supplement 38): 149.

Schoenemann, P. T., Gee, J., Avants, B., Holloway, R. L., Monge, J. & Lewis, J. (2007a). Validation of Plaster Endocast Morphology Through 3D CT Image Analysis. *American Journal of Physical Anthropology* 132: 183-192.

Schoenemann, P. T., Monge, J., Avants, B. B., Glotzer, D. & Gee, J. C. (2007b). Sex differences in cranial form assessed via non-rigid deformation analysis of high-resolution CT images. *American Journal of Physical Anthropology* 132 (Supplement 44): 209.

Schoenemann, P. T., Holloway, R. L., Avants, B. B. & Gee, J. C. (2008). Endocast asymmetry in pongids assessed via non-rigid deformation analysis of high-resolution CT images. *American Journal of Physical Anthropology* 135 (Supplement 46): 187-188.

Schoenemann, P. T., Monge, J., Glotzer, L. D. & Campana, M. (2008). The open research CT scan archive. *British Institute of Radiology Newsletter Spring*: 13-15.

Schoenemann, P. T., Monge, J., Avants, B. B. & Gee, J. C. (2009). An atlas of modern human cranial morphology constructed via non-rigid deformation analysis of high-resolution CT images. *American Journal of Physical Anthropology* 138 (Supplement 48): 231.

Schoenemann, P. T., Monge, J., Avants, B. B. & Gee, J. C. (2010). Creating statistical atlases of modern primate endocranial morphology using non-rigid deformation analysis of high-resolution CT images. *American Journal of Physical Anthropology* 141 (Supplement 50): 208-209.

Schoenemann, P. T., Holloway, R. L., Monge, J., Avants, B. B. & Gee, J. C. (2011). Differences in endocranial shape between Homo and Pongids assessed through non-rigid deformation analysis of high-resolution CT images. *American Journal of Physical Anthropology* 144 (Supplement 52): 265-266.

Websites cited

http://plum.museum.upenn.edu/~orsa/Welcome.html