

Artificial Neural Networks and Three-Dimensional Digital Morphology: A Pilot Study

Roger L. King^a A.L. Rosenberger^b L. Leann Kanda^c

^a Mississippi State University, Mississippi State, Miss., ^b Brooklyn College, CUNY, New York, N.Y., and ^c University of Massachusetts, Amherst, Mass., USA

Key Words

Artificial neural networks • Unsupervised learning • Laser digitizing • Three-dimensional digital morphology • Dentition • Anthropoids • Systematics

Abstract

In this pilot study, we used an unsupervised learning algorithm for self-organization and pattern matching to create feature maps that can be applied to morphological problems. We designed a network to analyze 83 first and/or second upper and lower molar sets representing 13 anthropoid primate species, based on three-dimensional measures obtained from laser-digitized, virtual specimens. As shown in a comparison with a principal-component analysis of the virtual specimens, the artificial neural network approach provided more biologically meaningful information than the conventional multivariate analysis approach. The methodology discovered partitions and hierarchical clusters consistent with anthropoid systematics, from the species (or subspecies) level to the highest categories, by sorting and allocating upper and lower molar teeth. As one might expect, measures of upper molars were richer in phenetic information than those of lower molars, even among the anatomically diverse platyrrhines. We also show that reducing taxonomic noise (i.e. biological variation) by limiting the analysis to a monophyletic subset improves discrimination.

Copyright © 2005 S. Karger AG, Basel

Introduction

In this report, we present a pilot study that applies an artificial neural network (ANN) architecture to a database of three-dimensional (3-D) digital morphology measures describing a diverse sample of anthropoid molar teeth. 3-D digital mor-

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

©2005 S. Karger AG, Basel

Accessible online at:
www.karger.com/fpr

Roger L. King, Department of Electrical and Computer Engineering, Box 9571, Mississippi State University Mississippi State, MS 39762-9571 (USA)
Tel. +1 662 325 2189, Fax +1 662 325 2298
E-Mail rking@ece.msstate.edu

phology, the study of form based upon 3-D digitized or virtual specimens, has many advantages [Ungar and M'Kirera, 2003], both theoretical and practical. The key benefits we exploit here are the potential to define and extract precise measures that are more realistic descriptors of complex anatomical form than caliper measures. Using accurately digitized specimens, one can take linear (point to point), conventional two-dimensional (2-D; compound point to point measures describing a plane) and 3-D measures, i.e. explicit multi-point dimensions of contours, surfaces and volumes which accurately represent structural shape or space, as defined by their geometric coordinates.

ANNs are hardware and software implementations based on biological models of circuitry and process used by humans to accomplish many fundamental tasks, such as pattern recognition, memory and visualization. Like the brain, ANNs employ simple processors with a high degree of parallelism. The strength of the ANN architecture used in this study that appeals to comparative morphology and systematics is its potential for knowledge discovery, pattern recognition, adaptability and self-organization.

Our pilot study has two goals: (1) to apply the approach of ANNs to a sample of anthropoid molar teeth in order to investigate its potential as a systematics tool and (2) to begin to evaluate the power of geometrically true 3-D measures, as opposed to standard caliper measures, as quantitative descriptors of molar morphology. The samples we have chosen do not reflect any particular systematics or anatomical problem. Rather, they were selected to represent a range of morphologies, a diversity of phylogenetic groups, a mixture of interspecific and intraspecific pairings of interest, and because we had convenient access to the material.

ANNs and Unsupervised Learning

ANNs are systems of simple processors (neurons) that are interconnected in an organized fashion (architecture). There are numeric values associated with the interconnections (weights) that are adjusted over time to create a mapping between a n-dimensional input space and a m-dimensional output space. These weights represent the knowledge about the problem domain. Changes in the weights are analogous to learning. A specific architecture will have a learning algorithm associated with it, and each of the neurons in the architecture will make use of a function (linear or nonlinear) to describe the neuron's internal state (activation) resulting from received stimuli. The architectures (neurons and their interconnections) are computational approaches simulating the biological neural network. Therefore, many of the architectures, including the one used in this study, are based on findings from the field of neuroscience [Kohonen, 1989].

Learning in an ANN can occur in either a supervised or an unsupervised fashion. A supervised approach uses a learning algorithm that creates an input/output mapping based on a labeled training set (i.e. an input vector is paired with a correct output vector). In this case, the network will learn a functional approximation from the input/output pairings and will have the ability to recognize or classify a new input vector into a correct output vector (generalization). An unsupervised learning architecture, in contrast, presents the network with only a set of unlabeled input vectors. Unsupervised learning is used for data compression, feature discovery and classification. The sole information available to the network is the correlations it derives from the input vectors. The network is expected to create charac-

terizations about the input vectors from these correlations and to produce outputs corresponding to a learned characterization (i.e. knowledge discovery) [King, 1998].

As with humans, learning with ANNs takes time. When using either a supervised or an unsupervised learning algorithm, data are presented to the network numerous times. One complete presentation of a set of training data is referred to as an epoch. As the data are repeatedly presented to the network, the learning algorithm evolves the weights connecting the neurons (i.e. learns). It is within the weights that knowledge is stored (represented) about the relationships being learned or discovered by the network.

ANNs that use unsupervised learning have the goal of determining natural clusters or feature similarity within the data set and to display its results in a meaningful manner. Since no labeled training sets are used in this approach, the outputs from the unsupervised learning network must be examined by a domain expert to determine if the classification provides any new insight into the data set. If the result is not reasonable, then an adjustment is made to one of the training parameters used to guide the network's learning, and the network is presented the patterns again. However, before examining the more complex self-organizing map (SOM) architecture, it is beneficial to recall a simpler form of unsupervised learning that makes use of a clustering algorithm. These networks utilize input data in a vector form, and a distance metric (e.g. Hamming distance, Euclidean distance) is used to determine nearness (similarity) of a new input pattern to an existing cluster. Decision options for the network are to include the new pattern into an existing cluster or to create a new cluster for this pattern. Learning (i.e. weight adjustment) normally occurs after a pattern has been assigned to a cluster. In the case of a clustering algorithm, this is simply adjusting the weights defining the center of the cluster. An example of this type of unsupervised learning network is k-means clustering [Moon and Stirling, 2000].

Self-Organizing Maps

Kohonen [1989] proposed that there is a general tendency in human information processing to compress complex information by forming reduced representations of the most relevant facts. An important aspect of this reduction in dimensionality is the ability to preserve the structure (interrelationships) of the information. He proposed that this simplified representation is accomplished via spatial ordering of processing units (neurons) within the brain. However, this ordering did not involve movement of any neurons but a change in the internal parameters of identical units.

A self-organizing feature map algorithm is used to convert patterns of arbitrary dimensionality into the responses of 2- or 3-D arrays of neurons (maps). A SOM may be thought of as a self-organizing cluster. It is clustering because of its tendency for data compression and it is self-organizing because similar clusters are spatially near on the map. The basic components of a 2-D SOM are shown in figure 1. Note that each element of the input vector x (description of unknown specimen) is connected to each of the processing units on the map through the weight vector w_{ij} . After training, the SOM will define a mapping between the input data space and the 2-D map of neurons. The output y_i of a processing unit is then a function of the similarity between the input vector and the weight vector. The nonlinear mapping

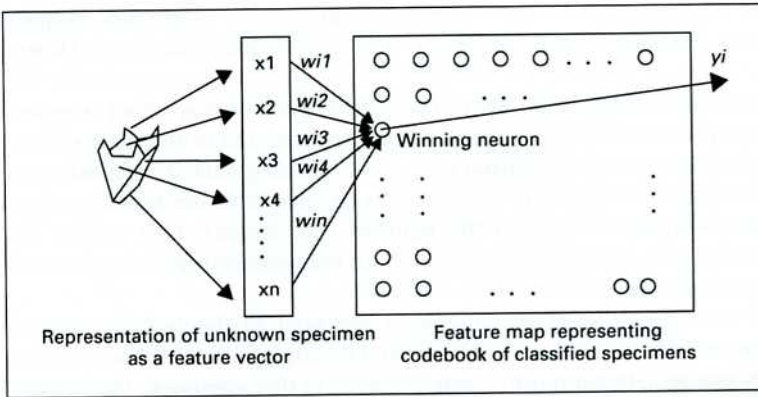


Fig. 1. Graphic depiction of data analysis using an ANN and a SOM (see text).

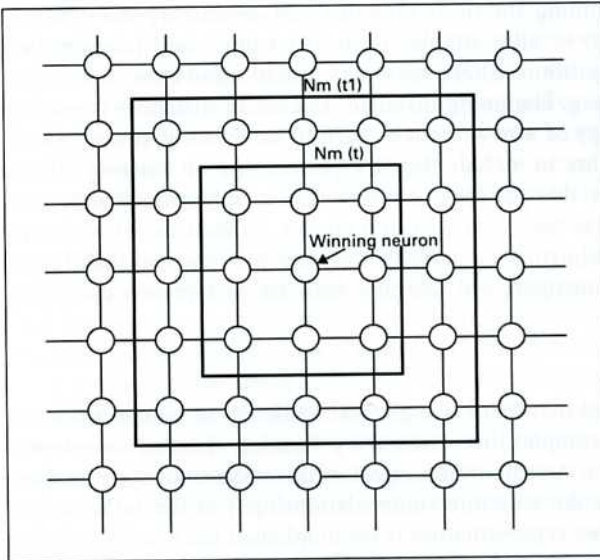


Fig. 2. Rectangular grid used for SOMs showing 4 nearest neighbors.

of the SOM utilizes a technique developed by Sammon [1969] that preserves the higher dimensional nearness on the map. In other words, if two vectors are near each other in high dimensional space, then they are near on the map.

Figure 1 shows how a trained feature map would respond with a winning output neuron when excited by an original training pattern or an unknown input vector pattern that is similar. Knowledge about the significance of the area around the winning neuron will then help the user in knowledge discovery. The SOM feature

map also functions as a nonlinear projection of the probability density function of the high dimensional input vector onto the 2-D display [Kohonen et al., 1996]. This means that following the formation of the self-organized internal representation for existing implicit relationships between the 12-D original training data (i.e. after learning), the same implicit input spatial relationships become explicit on the feature map. In simpler terms, if two points are close on the 2-D map, then they were also close in the higher dimensional space. In other words, after training, the internal representation for knowledge about the relationships within the original training data becomes explicit on the feature map.

The SOM algorithm is typically implemented on a planar array of neurons (fig. 2) with spatially defined neighborhoods (e.g. hexagonal or rectangular arrays, with 6 or 4 nearest neighborhoods, respectively). Also, the map must contain some method of compressing the data into a manageable form. One important attribute of a SOM is that it performs data compression without information being lost with regard to the relative distance between data vectors. A SOM typically uses the Euclidean distance to determine the relative nearness or similarity of data.

The idea of a spatial neighborhood, N_m , is used in measuring the similarity between the input vector and values of the reference vector represented by the vector of weights between the input layer and all of the neurons on the map. Before training begins, the weights are randomized and a learning rate and neighborhood size are selected. Then, when a training vector is presented to the network, it finds the neuron on the map with the most similar weight values. The weights of the winning neuron and the neighborhood neurons are then adjusted (learning) to bring them closer to the training vector. Over the course of the iterative training process, the neighborhood size and learning rate are independently decreased until the map no longer makes significant adjustments. The result is that the neurons within the currently winning neighborhood undergo adaptation at the current learning step while the weights in the other neighborhoods remain unaffected. The winning neighborhood is defined as the one located around the best matching neuron, m .

The operation of the SOM algorithm progresses as follows. First, for every neuron i on the map, there is associated a parametric reference vector w_i . The initial values of w_i (0) are randomly assigned. Next, an input vector x (R^n) is applied simultaneously to all of the neurons. The smallest of the Euclidean distances is used to define the best-matching neuron; however, other distance metrics may be explored to determine their efficacy in clustering the codebook vectors [King et al., 1993]. As the training progresses, the radius of N_m decreases such that $N_{m(t_1)} > N_{m(t_2)} > N_{m(t_3)} \dots$, where $t_1 < t_2 < t_3 \dots$. In other words, the neighborhood of influence can be very large when learning begins, but towards the end of learning, the neighborhood may involve only the winning neuron. The SOM algorithm also uses a learning rate that decreases with time.

In summary, the self-organization of the map proceeds as follows: (1) the map is presented with a sufficient number of training patterns; (2) weights are only adjusted on the neurons in the winning neighborhood, and (3) the adjustment is in proportion to the activation received by each neuron in the neighborhood. The effect of this weight adjustment rule is a tendency to enhance the same responses to a sufficiently similar subsequent input. As a result, a map is obtained with weights coding the stationary probability density function of the pattern vectors used for the

training. The map also displays the data from a different viewpoint; instead of viewing the data as a n-dimensional vector, it can be viewed as a 2-D plot. This is where expert human analysis is enhanced. Instead of looking at the n-dimensional input vector of a sample and trying to determine what its meaning is, one need only look at the location of the sample on the map.

Materials and Methods

Measurements

We used a high-resolution laser scanner with an accuracy of 0.025 mm to generate an array of x,y,z coordinate points to describe 83 first and/or second upper and lower molar teeth of 13 anthropoid primate species. Each of the specimens was digitized at a density of approximately 150 points/mm². This density level faithfully captures surface detail in very small specimens. Using this figure as a standard also means that the amount of surface relief described per unit area is comparable across all cases. This serves to minimize bias in measurements of area and volume, which would be sensitive to the size and shape (topographic) differences present in the sample. Our methodology for collecting coordinate data and the editing process involved in deriving measurements from the data – such as the demarcation of anatomical features for measurement – has been proven in numerous repeat trials [Rosenberger and Kellogg, unpubl. observations]. Negligible error was found in measurement replication tests within and between specimens. Linear measures taken on population samples of virtual specimens were also statistically indistinguishable from typical caliper measurements taken on the actual specimens.

The following measurements were used:

Laser basal length (LBL): on a digital specimen aligned with the mesiodistal axis parallel to the x -axis, LBL is the distance between minimum and maximum x -coordinate values in the x -axis.

Laser basal breadth (LBB): on a digital specimen aligned with the mesiodistal axis parallel to the x -axis, LBB is the distance between minimum and maximum y -coordinate values found anywhere between the occlusal table of the crown and the cervix.

Laser basal area: $LBL \times LBB$ (mm²).

Occlusal table area: area of the occlusal table, or 'chewing surface' of upper and lower molars, i.e. the aspect contained within the boundaries of the cusp-crest system and excluding the crown sidewalls, except for the wear facets of the upper lingual notch and lower buccal ectoflexid.

Cusp height: height of a cusp tip above the lowest point in the trigon or talonid basin of an upper or lower molar, respectively, in the z -axis.

Apical distance: the oblique distance between the apical tip of a cusp and the lowest point in the trigon or talonid basin of an upper or lower molar, respectively.

Samples

Species comprising the 83 cases used in this study are listed in table 1. The living forms range in weight, approximately, from a 300 g *Callithrix penicillata* to an 11,000 g *Colobus satanus*; molar lengths range from 2.3 to 7 mm, respectively, to 10 mm in the fossil catarrhine *Oreopithecus bambolii*. We digitized sharp epoxy or plaster casts of molars or tooth rows showing little or no detectable wear rather than actual specimens; mounting jaws would have required additional setup time. Except for a few fossils, each individual was represented by an occluding set of teeth. In the extant species, except for callitrichines, an upper and lower second molar set was used. First molars were substituted for the M2s in callitrichines, as the second molars are highly reduced even in the three-molared *Callimico*. For the fossils, these conditions were relaxed. All have been identified as likely second molars, but uppers and lowers did not necessarily belong to the same individuals. In a few cases, including *Oreopithecus*, the fossil samples were also moderately worn.

Table 1. Samples and measurements used in this study

Taxon	Pat- tern No.	Hypo- conid Z	Hypo- conid AD	Proto- conid Z	Proto- conid AD	Meta- conid Z	Meta- conid AD	Ento- conid Z	Ento- conid AD	LBL	LBB	LBA	OTA
<i>Colobus satanas</i>	1	1.58	3.27	1.62	2.96	2.92	3.83	2.29	3.13	6.75	6.32	42.66	43.51
<i>Ptilocolobus pennantii</i>	2	1.31	2.40	1.45	2.38	2.14	2.47	1.87	2.28	6.25	4.14	25.88	24.12
<i>Procolobus verus</i>	3	1.85	3.67	1.84	2.99	2.14	2.78	1.70	2.49	5.13	5.04	25.86	31.65
<i>Macaca nigra</i>	4	1.36	2.40	1.69	3.20	2.75	4.42	2.20	3.11	8.37	7.78	65.12	63.49
<i>Oreopithecus bambolii</i>	5	0.19	2.53	0.38	3.39	1.23	3.22	0.34	3.49	9.88	7.58	74.89	85.89
<i>Oreopithecus bambolii</i>	6	0.71	2.77	1.41	3.83	1.34	3.44	0.40	3.40	9.50	7.75	73.63	65.10
? <i>Prohylobates</i> sp.	8	2.37	3.03	1.67	2.30	1.72	2.88	2.10	3.08	6.75	6.22	41.99	27.42
? <i>Prohylobates</i> sp.	9	1.35	3.26	1.82	3.75	2.52	3.37	1.35	2.06	8.12	7.12	57.81	55.32
<i>Aotus trivirgatus</i>	10	0.57	1.28	0.65	1.57	1.23	1.66	0.95	1.45	3.50	3.07	10.75	8.99
<i>Callicebus moloch</i>	11	0.58	1.18	0.97	2.29	1.65	2.46	1.01	1.70	3.37	3.20	10.78	10.51
<i>Callicebus moloch</i>	12	0.74	1.35	1.27	2.54	1.56	2.14	0.76	1.25	3.38	3.09	10.44	10.13
<i>Callicebus moloch</i>	13	0.79	1.39	1.13	2.46	1.18	1.97	0.59	1.13	3.62	3.03	10.97	10.88
<i>Saimiri sciureus</i>	14	0.52	1.27	0.82	2.45	1.38	2.17	0.82	1.08	3.75	3.01	11.29	10.40
<i>Saimiri sciureus</i>	15	0.73	1.23	1.03	1.62	1.08	1.47	0.56	1.01	2.75	2.66	7.32	6.91
<i>Saimiri sciureus</i>	16	0.69	1.21	1.17	1.67	1.17	1.50	0.50	1.07	2.61	2.70	7.05	6.39
<i>Saimiri sciureus</i>	17	0.73	1.14	1.05	1.86	1.10	1.61	0.62	0.90	2.88	2.67	7.69	6.86
<i>Saimiri sciureus</i>	18	0.61	1.15	1.13	1.69	1.28	1.68	0.59	1.07	2.62	2.63	6.89	6.55
<i>Saimiri sciureus</i>	19	0.84	1.34	0.97	1.66	1.23	1.50	0.90	1.11	2.63	2.56	6.73	6.44
<i>Saimiri sciureus</i>	20	0.74	1.28	0.91	1.91	0.84	1.34	0.54	0.95	2.75	2.85	7.84	5.97
<i>Saimiri sciureus</i>	21	0.66	1.22	1.04	1.78	1.26	1.66	0.76	1.02	2.50	2.60	6.50	6.30
<i>Saimiri sciureus</i>	22	0.75	1.16	1.03	1.40	1.22	1.68	0.56	1.11	2.63	2.78	7.31	6.04
<i>Saimiri sciureus</i>	23	0.78	1.33	1.06	1.56	1.26	1.60	0.68	1.09	2.50	2.53	6.33	6.06
<i>Saimiri sciureus</i>	24	0.75	1.22	0.92	1.37	1.07	1.54	0.57	1.19	2.63	2.66	7.00	6.32
<i>Callithrix jacchus</i>	25	0.50	1.02	0.82	1.37	1.00	1.36	0.38	0.66	2.12	2.05	4.35	4.06
<i>Callithrix jacchus</i>	26	0.58	1.03	0.99	1.60	0.96	1.34	0.48	0.73	2.00	2.13	4.26	3.93
<i>Callithrix jacchus</i>	27	0.68	0.91	0.93	1.44	0.76	1.18	0.33	0.62	2.50	1.99	4.98	4.12
<i>Callithrix jacchus</i>	28	0.81	0.97	1.03	1.64	0.84	1.24	0.38	0.73	2.25	2.14	4.82	4.30
<i>Callithrix jacchus</i>	29	0.54	0.82	0.81	1.32	0.80	1.31	0.28	0.74	2.13	2.26	4.81	3.82
<i>Callithrix jacchus</i>	30	0.67	1.09	0.96	1.53	0.76	1.22	0.35	0.54	2.12	2.05	4.35	4.28
<i>Callithrix penicillata</i>	31	0.64	1.02	1.01	1.45	1.06	1.41	0.56	0.97	2.25	2.27	5.11	3.99
<i>Callithrix penicillata</i>	32	0.60	0.83	0.90	1.50	0.98	1.33	0.48	0.89	2.50	2.12	5.30	4.54
<i>Callithrix penicillata</i>	33	0.80	0.99	1.05	1.41	1.07	1.32	0.53	0.88	2.25	2.19	4.93	4.68
<i>Callithrix penicillata</i>	34	0.62	0.95	0.91	1.35	0.99	1.25	0.43	0.79	2.13	2.13	4.54	4.46
<i>Callithrix penicillata</i>	35	0.83	1.09	1.02	1.66	1.00	1.39	0.49	0.86	2.38	2.10	5.00	4.52
<i>Callimico goeldii</i>	36	0.74	1.25	0.92	1.66	0.75	1.06	0.75	1.25	3.00	2.46	7.38	6.86
<i>Callimico goeldii</i>	37	1.05	1.25	1.14	1.63	1.14	1.57	0.65	1.06	2.62	2.41	6.31	6.27
<i>Callimico goeldii</i>	38	0.72	1.12	0.96	1.42	1.03	1.34	0.42	0.96	2.38	2.33	5.55	5.28
<i>Callimico goeldii</i>	39	0.53	0.83	0.84	1.39	1.32	1.65	0.56	1.05	2.37	2.39	5.66	5.57
<i>Callimico goeldii</i>	40	0.92	1.48	1.21	1.46	1.27	1.51	0.41	1.07	2.62	2.31	6.05	5.95
<i>Colobus satanas</i>	41	2.03	3.27	2.17	3.63	2.12	3.17	1.99	2.77	6.63	7.58	50.26	48.89
<i>Ptilocolobus pennantii</i>	42	0.56	3.32	0.64	4.06	2.61	3.48	1.68	2.34	6.50	6.81	44.27	45.97
<i>Ptilocolobus pennantii</i>	43	1.36	2.59	1.67	2.51	2.08	2.41	1.69	2.15	6.12	5.23	32.01	24.72
<i>Procolobus verus</i>	44	1.01	2.30	1.31	3.15	2.07	2.42	1.36	2.59	5.13	5.33	27.34	32.30
<i>Macaca nigra</i>	45	1.61	2.90	1.91	3.86	1.75	3.29	1.63	2.42	8.75	8.99	78.66	57.79
<i>Oreopithecus bambolii</i>	46	0.81	3.33	0.83	3.27	1.48	4.38	0.75	3.25	11	9.74	107.14	77.73
<i>Oreopithecus bambolii</i>	47	0.66	2.72	0.69	2.82	0.91	3.16	0.37	1.56	8.98	9.04	81.18	57.44
<i>Oreopithecus bambolii</i>	48	0.59	3.24	1.06	2.96	0.83	3.99	0.30	2.98	9.37	9.63	90.23	69.31
<i>Rangwapithecus gordoni</i>	49	1.40	5.77	1.87	4.69	1.56	3.02	1.18	2.33	8.75	9.82	85.93	72.67
? <i>Prohylobates</i> sp.	50	1.68	2.89	1.53	2.72	1.80	2.87	1.57	2.41	7.50	8.95	67.13	46.21
? <i>Prohylobates</i> sp.	51	1.32	3.29	1.22	3.33	1.17	2.21	1.28	1.97	7.00	8.45	59.15	34.41
? <i>Prohylobates</i> sp.	52	1.75	2.49	1.25	2.48	1.34	2.49	1.48	2.12	7.38	8.23	60.74	37.14
<i>Aotus trivirgatus</i>	53	0.95	1.45	1.23	1.66	0.65	1.57	0.57	1.28	3.12	3.97	12.39	8.89

Table 1 (continued)

Taxon	Pat- tern No.	Hypo- conid Z	Hypo- conid AD	Proto- conid Z	Proto- conid AD	Meta- conid Z	Meta- conid AD	Ento- conid Z	Ento- conid AD	LBL	LBB	LBA	OTA
<i>Aotus trivirgatus</i>	54	0.90	1.92	0.95	1.76	0.98	1.68	1.06	1.60	3.37	3.99	13.45	10.67
<i>Callicebus moloch</i>	55	0.92	2.22	1.08	1.96	1.06	1.55	0.86	1.35	3.37	4.36	14.69	9.77
<i>Callicebus moloch</i>	56	0.83	1.71	1.00	1.58	1.13	1.74	1.05	1.62	3.37	4.16	14.02	9.58
<i>Callicebus moloch</i>	57	0.63	2.44	0.76	1.67	1.02	1.40	0.85	1.57	3.25	4.26	13.85	9.02
<i>Saimiri sciureus</i>	58	0.67	1.87	0.96	1.44	1.00	1.56	0.58	1.15	2.63	3.69	9.70	4.96
<i>Saimiri sciureus</i>	59	0.62	1.81	0.82	1.68	1.19	1.67	0.84	1.04	2.5	3.85	9.63	5.68
<i>Saimiri sciureus</i>	60	0.67	1.76	1.03	1.37	1.12	1.40	0.75	1.28	2.5	3.90	9.75	5.15
<i>Saimiri sciureus</i>	61	0.70	1.56	1.09	1.46	1.13	1.60	0.59	1.06	2.38	3.83	9.12	5.31
<i>Saimiri sciureus</i>	62	0.68	1.59	1.15	1.48	1.22	1.54	0.66	1.00	2.25	3.53	7.94	4.99
<i>Saimiri sciureus</i>	63	0.45	1.86	0.80	1.22	1.09	1.58	0.70	1.21	2.5	4.03	10.08	5.71
<i>Saimiri sciureus</i>	64	0.62	1.38	1.08	1.29	1.05	1.74	0.54	0.94	2.5	3.75	9.38	5.68
<i>Saimiri sciureus</i>	65	0.84	1.85	1.37	1.71	1.15	1.64	0.59	1.08	2.38	3.92	9.33	6.40
<i>Saimiri sciureus</i>	66	0.57	1.67	1.00	1.35	1.14	1.53	0.69	1.08	2.5	3.75	9.38	6.19
<i>Saimiri sciureus</i>	67	0.65	1.49	1.12	1.28	1.39	1.72	0.61	1.33	2.5	3.89	9.73	5.69
<i>Callithrix jacchus</i>	68	0.00	0.00	0.64	1.07	0.73	1.16	0.94	1.46	2.38	2.92	6.94	3.92
<i>Callithrix jacchus</i>	69	0.00	0.00	0.44	0.93	0.97	1.42	0.87	1.30	2.37	3.00	7.11	3.73
<i>Callithrix jacchus</i>	70	0.00	0.00	0.68	0.98	0.76	1.38	0.68	1.01	2.12	2.87	6.08	3.47
<i>Callithrix jacchus</i>	71	0.00	0.00	0.86	1.08	0.89	1.22	0.91	1.49	2.25	3.24	7.29	3.91
<i>Callithrix jacchus</i>	72	0.00	0.00	0.61	0.99	0.92	1.39	0.79	1.10	2.13	3.05	6.50	2.94
<i>Callithrix jacchus</i>	73	0.00	0.00	0.62	0.96	0.95	1.16	0.74	1.32	2.25	3.03	6.82	3.49
<i>Callithrix penicillata</i>	74	0.00	0.00	0.74	1.01	0.80	1.35	1.06	1.52	2.38	3.05	7.26	3.59
<i>Callithrix penicillata</i>	75	0.00	0.00	0.71	0.89	0.71	1.17	0.77	1.34	2.37	3.00	7.11	3.00
<i>Callithrix penicillata</i>	76	0.00	0.00	0.85	0.99	0.72	1.17	0.89	1.52	2.12	3.07	6.51	3.20
<i>Callithrix penicillata</i>	77	0.00	0.00	0.83	1.02	0.78	1.28	0.88	1.41	2.25	3.01	6.77	3.57
<i>Callithrix penicillata</i>	78	0.00	0.00	0.91	1.05	0.84	1.18	0.79	1.24	2.25	2.91	6.55	3.61
<i>Callimico goeldii</i>	79	0.00	0.00	0.95	1.39	1.66	1.90	0.87	1.31	2.50	3.52	8.80	4.78
<i>Callimico goeldii</i>	80	0.00	0.00	0.99	1.52	1.25	1.70	0.93	1.13	2.34	3.73	8.73	4.14
<i>Callimico goeldii</i>	81	0.00	0.00	0.91	1.30	1.25	1.72	0.91	1.42	2.38	3.41	8.12	3.92
<i>Callimico goeldii</i>	82	0.00	0.00	0.89	1.24	1.28	1.62	0.61	1.11	2.12	3.70	7.84	3.12
<i>Callimico goeldii</i>	83	0.00	0.00	0.83	1.45	1.47	1.86	0.73	1.17	2.37	3.51	8.32	4.98

Individual specimens are represented by sequential numbers as a pattern. Measurements are expressed in millimeters and square millimeters, respectively. Measurements of each cusp include the vertical height (Z) and oblique apical distance (AD) from the cusp tip to the lowest point of the trigonid or talonid basin. LBL = Laser basal length; LBB = laser basal breadth; LBA = laser basal area; OTA = occlusal table area.

Procedures

After scanning, each of the specimens was visualized on-screen to verify the integrity of the digitization process. They required minimal editing to delete stray points. Obvious measurement error (e.g. due to spectral reflection) and extraneous information that was captured by the digitizer, such as the mounting material (clay or hot glue) used to stabilize the specimen during the scanning process, was removed. For this study, we also clipped out of the virtual specimens the occlusal table of each tooth in order to measure the occlusal table area. All in all, postprocessing took no more than 15 min per specimen.

The Stuttgart Neural Network Simulator (SNNS) was the software environment selected for this research. The SNNS software was developed by the University of Stuttgart's Institute for Parallel and Distributed High-Performance Systems [1998]. The goal was to create an efficient and flexible simulation environment for research on and application of neural networks. The software is presently being maintained by the University of Tübingen and is available for download at www-ra.informatik.uni-tuebingen.de/SNNS/.

As explained above, learning within the SOM is time dependent in two ways. First, the neighborhood size decreases with time, and, second, the amount of learning within the winning neighborhood also decreases with time. Therefore, the learning parameters to be specified by the user include the size of the map, initial learning rate and neighborhood size, rates at which these parameters decrease and number of iterations (epochs). The final design parameters used in this pilot study were determined after several experiments had been conducted to gain insight from variations of these parameters on final map outputs. For example, a 10×10 map size was the final selection for a map size to accommodate all 83 virtual specimens. However, smaller map sizes were used (e.g. an 8×8 map size) but an analysis of the maps resulting from these tests was deemed of less interest because it forced the network to severely compress the data set.

Each of the 83 cases was stripped of their identifying information for the analyses so the ANN was, in essence, performing a blind classification test. The order in which each case was entered into the ANN was also randomized. The final design of all the SOMs used in the study consisted of the following parameters: a rectangular neighborhood; horizontal size, 10 neurons; vertical size, 10 neurons; initial neighborhood size, 6; epochs, 3,000. In other words, each specimen was allowed 3,000 iterations to localize in a 10×10 SOM by pattern matching against surrounding cases of a diminishing neighborhood size, beginning with a radial distance of 6 cells away from the new target specimen. The weights that generated our final SOM are provided in Appendix 1.

To help visualize the results of the SOM and compare them against other ways of treating the measurement data, we used minimum spanning trees (MSTs). The advantage of a MST in this context is that it arranges the information into a hierarchical clustering of cases, whereas our SOMs display cases in a 2-D array of adjacencies. A spanning tree is an undirected graph (V, E) that is connected and acyclic. It consists of some integer number of vertices (V) and edges (E) connecting the vertices. In the case of the 10×10 SOM, we have 100 vertices and edges of $O(V^2)$ or 10,000. The MST is the spanning tree on the vertices that has the minimal total length apportioned to its edges. In our case, an edge corresponds to the Euclidean distance between two vertices. The MST algorithm used in this study was adopted from Papadimitriou and Steiglitz [1982]. To evaluate and help visualize the discriminatory effects of an ANN based upon digital morphology measures, we also present MSTs based on raw descriptive data, of caliper and 3-D measures.

Results

Background Analysis

Figure 3 is a MST based on three conventional metrics, molar length (LBL), breadth (LBB) and crown cross-sectional area (LBL \times LBB). As expected, the array is mainly influenced by size, as evidenced by its linearity and lack of parallel adjacencies. Specimens having the smallest measurements, the callitrichines, form one stem of the MST opposite those having the largest dimensions, the catarrhines. Within platyrrhines, most lower and upper molars are segregated by size. This has the effect of sorting some species as well. The uppers and lowers of *Callicebus* and *Aotus* are not well separated from one another, however. Elsewhere among platyrrhines, the only notable branching structure in the MST is confined to its central section, representing upper molars of *Callithrix* species, *Callimico* and *Saimiri*. The importance of this, like the few small branches evident elsewhere, is moot. Catarrhine molars present no clear upper-lower molar separation because of the near mirror image construction of their bilophodont cheek teeth. This same effect is repeated throughout the analyses. The limited sorting of the catarrhine species that does occur is also an effect of size differences.

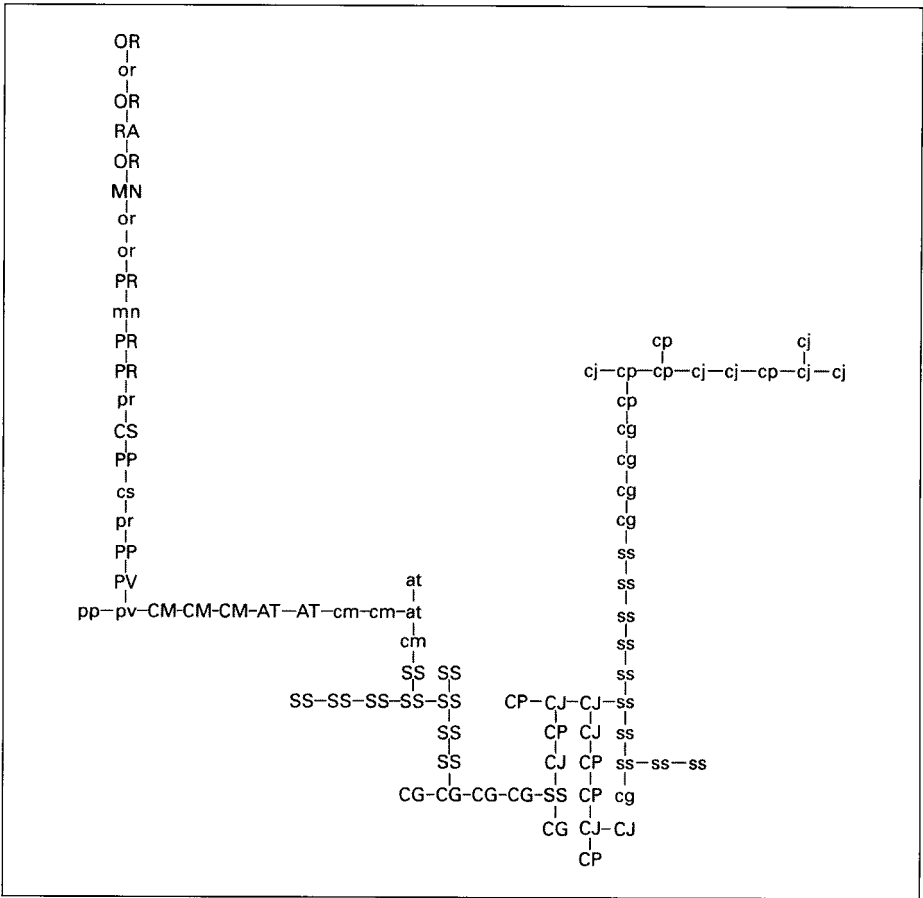


Fig. 3. MST based on untransformed linear measures. Upper case and lower case for all figures denote upper and lower molars, respectively. Letter abbreviations for all figures: Platyrrhines – CJ = *Callithrix jacchus*; CP = *Callithrix penicillata*; CG = *Callimico goeldii*; SS = *Saimiri sciureus*; CM = *Callicebus moloch*; AT = *Aotus trivirgatus*; Catarrhines – CS = *Colobus satanus*; MN = *Macaca nigra*; PR = *?Prohylobates* sp.; OR = *Oreopithecus bambolii*; PP = *Ptilocolobus pennantii*; PV = *Procolobus verus*; RA = *Rangwapithecus gordonii*.

Using 3-D measures describing cusp height, cusp location and occlusal table area as descriptors alters the associative potential of the MST in a few ways, but it does not demonstrably improve the analysis for the entire sample (fig. 4). Overall, there is less linearity and more adjacency in the display because individuals of some species are clustered more tightly, ostensibly because of the richer input information. A size vector continues to dominate the output, but there is also some sorting based on shape and/or taxonomic effects. For example, *Callimico* uppers are no longer associated with *Saimiri* (fig. 3), and they form a coherent group linked with *Callithrix*. In addition, *Saimiri*, *Callicebus* and *Aotus* uppers are tightly

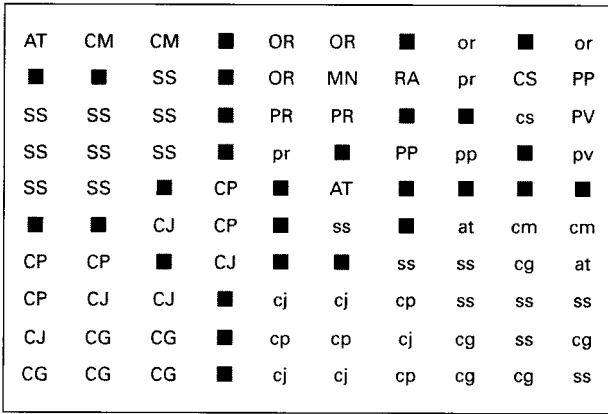


Fig. 5. SOM based on untransformed 3-D measures (see text).

lowers appear in the matrix. The other specimens have been suppressed as redundant cases.

Turning to the MST based on the weights of each of the cases that were classified in the SOM (fig. 5) as well as the blanks, the overall complexity of the tree, with numerous well-defined clusters and linkages, is profoundly different from the relatively linear MSTs based on raw data (fig. 3, 4). Groups are strongly partitioned or anchored by the empty cells of the SOM. Several clusters are rooted to the backbone of the MST through blanks. They depict the hierarchy of pattern matching decisions linking each of the individual cases. There are several interesting points of correspondence between this SOM tree and those based on raw measures. First, *Aotus* and *Callicebus*, together, continue to occupy adjacent space near the left central core of the MST. This result appears to be driven by resemblance in their upper molars and by mutual resemblances shared with *Saimiri* (in M^2) and the catarrhines (in $M^{2/2}$) as well. All of these forms have quadrate upper crowns. Second, the lower molars of individual platyrrhine species, although grouping together in a salient partition, are less well resolved taxonomically by the SOM than are their upper molars.

The three most coherent clusters defined by the MST (fig. 6) represent callitrichine upper molars, *Saimiri* uppers and all catarrhine teeth. An offset comprising a unit of *Saimiri* + *Callithrix* + *Callimico* uppers is evident, but this also includes an anomaly, the artificially segmented *Saimiri* group. This two-tier organization of cebid New World monkeys derived from the SOM data differs in interesting ways from the previous MSTs. In the background study of raw data, caliper measures linked *Saimiri* uppers and *Callimico* + *Callithrix* uppers (fig. 3), presumably because of size. Digital morphology measures associated *Saimiri* uppers with other four-cusped platyrrhines, *Aotus* and *Callicebus*, apparently based on form. The SOM-based MST spatially separates *Saimiri* uppers from the latter, but places them as a point of linkage to the three-cusped callitrichines. It also places *Saimiri* lowers well within the callitrichine cluster.

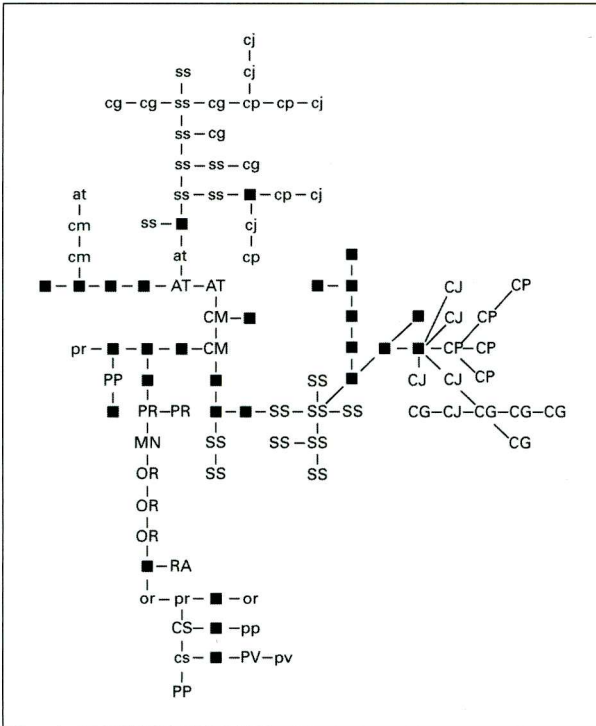


Fig. 6. MST based on SOM weights derived from 3-D measures.

An improved clustering is well illustrated among callitrichines. Uppers of the 5 *C. pennicillata* individuals are linked intraspecifically in a hierarchy, offset interspecifically through a blank space (implying absence of an existing case match), but clearly associated with the *C. jacchus* group. Four of 5 *C. jacchus* are also interlinked. However, there is no clear separation of *C. jacchus* from *Callimico* uppers due to the misclassification of the 1 *C. jacchus* outlier. Were this case removed, it seems likely all *Callimico* individuals would be identified as a discrete unit. If so, the aggregation of *C. jacchus* might also resolve. The well-defined *Saimiri* uppers are also split up, with 2 of 9 individuals appearing as outliers.

Among catarrhines, systematic improvements are less evident (fig. 6). There is a tendency to divide colobines from the others, as before (fig. 3, 4), and a few individual cases and taxa are better sorted out. The single *Rangwapithecus* upper molar and 2 *?Prohylobates* lowers, for example, are outliers as expected. One *?Prohylobates* tooth is fairly worn while the other is not. *Oreopithecus* uppers, and *Colobus satanus* uppers and lowers, also form tighter clusters. Overall, the main benefit here is that catarrhines are strongly demarcated from platyrrhines at the cluster root.

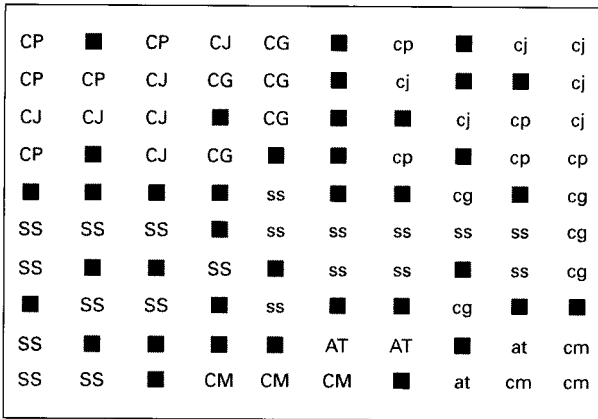


Fig. 7. SOM based on 3-D measures, platyrrhines only.

Revising the Data Set

In an effort to reduce taxonomic noise specific to our sample (which was constructed in an intentionally lax fashion, but with an eye toward diversity), we examined platyrrhines more closely by eliminating catarrhines from the analysis (i.e. reduced biological variation). A SOM based on the 62 New World monkey cases produced similar results in terms of clusters and within-group links, and a much simpler MST (fig. 7, 8). Several points are of interest. First, the 2-D matrix (fig. 7) sorted most platyrrhines by a tooth axis (horizontally) and a taxonomic axis (vertically), similar to the all-cases SOM: lower teeth on the right and upper teeth on the left of the matrix, with the exception of *Aotus* and *Callicebus*. Also, nearly all of the top 40 cells display *Callithrix* molars, uppers on the left and lowers on the right, strongly divided by a trace of blanks. The bottom 2 rows on the right are exclusively M2s of *Aotus* and *Callicebus*. There is a separation of *Saimiri* M2s in the middle and left lower quadrant of the SOM. Only the *Callimico* molars appear in cells widely offset diagonally in the SOM, with uppers and lowers falling far apart on the margins of the two *Callithrix* arrays.

In general, the MST (fig. 8) based on this SOM also sorts taxa from left to right, whereas the top and bottom clusters relate to upper and lower molars, respectively. This pattern is repeated in the central section of the backbone, where *Aotus* and *Callicebus* are represented. As with all the previous tests, here the uppers tend to give a better definition than the lowers. In addition, as with the SOM based on the full platyrrhine-catarrhine sample, *Saimiri* and callitrichine uppers are segregated, but linked to one another. *Saimiri* lowers, and *Callicebus* and *Aotus* uppers and lowers, are the connecting elements. *Saimiri* uppers are also less diffusely organized than in the platyrrhine-catarrhine sample, but the branching structure of the lowers is not easy to interpret. Nevertheless, the *Callithrix* molars are clustered together here, whereas they were divided into two groups in the larger sample. Another interspecific improvement appears here for the first time: *Aotus* and *Callicebus* upper molars are clearly sorted into 2 clusters.

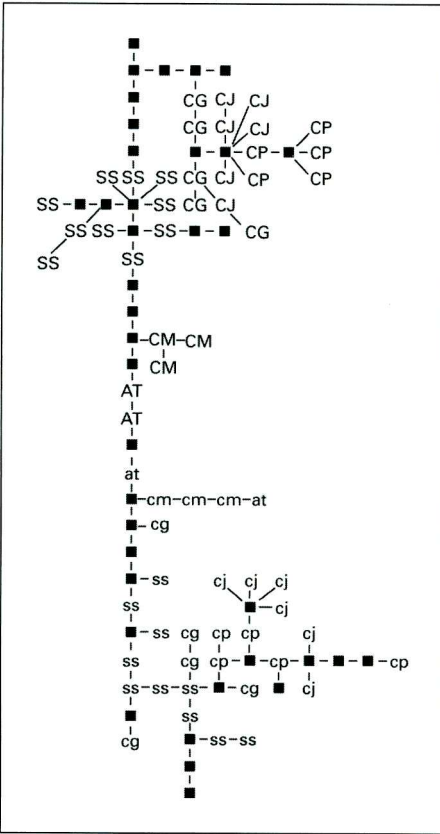


Fig. 8. MST based on SOM weights derived from 3-D measures, platyrrhines only.

SOM Comparison to a Principal-Component Analysis

We also conducted a principal-component analysis (PCA) of the 83 virtual specimens as a comparison with a more conventional multivariate analysis approach. PCA exploits the fact that in many cases where the dimension of the input vector is large (12-D in our case) the components of the vectors may be highly correlated (redundant). PCA is often used for its ability to orthogonalize the components of the input vector (so that they are uncorrelated with each other), to order the resulting orthogonal components (principal components) so that those with the largest variation come first and to eliminate those components that contribute the least to the variation in the data set. The first 5 principal components for the full sample of taxa based on 3-D measures are provided in Appendix 2.

Figure 9 and the data in Appendix 2 show that size evidently drives PCA 1, as expected when using untransformed measurements. Size was also a strong factor in the SOM, which distributes all the larger taxa in the right upper quadrant, with the smallest taxa in the left lower quadrant. One obvious result for the PCA is that, as a