Who Is LB1? Discriminant Analysis for the Classification of Specimens*

Aleix M. Martinez and Onur C. Hamsici
The Ohio State University, Columbus, OH 43210

April 28, 2008

Abstract

Many problems in paleontology reduce to finding those features that best discriminate among a set of classes. A clear example is the classification of new specimens. However, these classifications are generally challenging because the number of discriminant features and the number of samples are limited. This has been the fate of LB1, a new specimen found in the Liang Bua Cave of Flores. Several authors have attributed LB1 to a new species of *Homo*, *H. floresiensis*. According to this hypothesis, LB1 is either a member of the *early Homo* group or a descendent of an ancestor of the Asian *H. erectus*. Detractors have put forward an alternate hypothesis, which stipulates that LB1 is in fact a microcephalic modern human. In this paper, we show how we can employ a new Bayes optimal discriminant feature extraction technique to help resolve this type of issues. In this process, we present three types of experiments. First, we use this Bayes optimal discriminant technique to develop a model of morphological (shape) evolution from *Australopiths* to *H. sapiens*. LB1 fits perfectly in this model as a member of the *early Homo* group. Second, we build a classifier based on the available cranial and mandibular data appropriately normalized for size and volume. Again, LB1 is most similar to *early Homo*. Third, we build a brain endocast classifier to show that LB1 is not within the normal range of variation in *H. sapiens*. These results combined support the hypothesis of a very early shared ancestor for LB1 and *H. erectus*, and illustrate how discriminant analysis approaches can be successfully used to help classify newly discovered specimens.

Keywords: Pattern recognition, paleontology, discriminant analysis, morphological model, physical anthropology, Homo floresiensis.

1 Introduction

Two of the oldest questions to appeal humans are "who are we?" and "where do we come from?" The study of human evolution is at the core of these questions. This research is often guided by an analysis of available fossils. In studies in paleoanthropology (paleontology and physical anthropology), for example, bone measurements are commonly utilized for the classification of

*Pattern Recognition, *to appear*.

hominids within a set of categories. This involves the use of pattern recognition and shape analysis techniques to either construct classifiers [2] or build morphological models of evolution [26].

To date, researchers have mainly used Principal Component Analysis (PCA) [15] to address both issues – classification and evolutionary shape modeling. Unfortunately, PCA is only concerned with feature variance (and between-feature covariances, to be precise) and does not search for those linear combination of features that can best discriminate between classes. If one of the features in several of the classes carries a large intra-class variation, this will be selected by PCA as the most representative one. This may correspond to a feature that has a similar range of values in all species but a large variation within each species. Such a basis (feature) would be of little use to help us classify specimens. To resolve this problem, we can use supervised feature extraction algorithms, i.e., discriminant analysis [11]. In this case, we want to find that one or those two dimensions which project sample vectors of the same class close together while projecting those of distinct classes as far from each other as possible. The most popular of these algorithms is Linear Discriminant Analysis (LDA) [10]. However, to perform accurately, LDA requires that we extract $C-1$ features from the original space of d dimensions. Since one usually has four or more classes, LDA becomes impractical and far from optimal [21, 20]. Note that in paleoanthropology one is typically interested with that one or maybe two linear combinations of features that best discriminate classes. This is so because it is precisely such features which allow us to carry out further classifications of specimens accurately and because these dimensions specify which set of original features have been most active during evolution. These results can then be employed to construct evolutionary traits and trees, and can serve to determine behavior patterns [16].

A recently proposed algorithm [13] does recover the necessary optimal linear combination of features — optimal referring to maximizing classification not variance. Since classification is maximized when the Bayes error is minimized, the goal of [13] is to minimize the Bayes error. This is why we will refer to this method as Bayes optimal discriminant analysis. In the present paper, we show how one can use this algorithm to determine the classification of new specimens and to construct morphological models of shape change. The latter can then be used to extract the above mentioned important evolutionary conclusions needed to advance our knowledge of science.

Yet, this process is generally made difficult due to the limited amount of fossils available. We present a methodology to resolve this issue, which allows us to define morphological (shape) changes and low-dimensional discriminant spaces. As an illustration, we present a detail case study on LB1, a specimen originally classified as a new species of *Homo* — *H. floresiensis*.

1.1 H. floresiensis

In 2003 the skeleton of a small, not very old specimen was found in Liang Bua, in the small island of Flores (Indonesia) [22]. This specimen is referred to as LB1 [4]. LB1 seemed to correspond to a new (previously unknown) type of *Homo*. The original discovery team claimed LB1 may be related to *early Homo* [4, 23], while [9] analysis of the endocast suggested it may be related to *H. erectus* or an ancestor of these. More recently, [1] extended the original analysis and concluded that LB1 may be an archaic *Homo* or related to a yet to be discovered specimen which was in evolutionary transition between *Australopithecus* and *Homo*. Critics, however, argue that LB1 is nothing else than a diseased (microcephalic) *H. sapiens* [27, 18, 19, 14, 24].

To study this discrepancy and determine the most probable origins of LB1, one can conduct a statistical analysis of the data available. In such a study, a set of d bone measurements for a total of C classes (in this case, C species or genera) are interpreted as dimensions of a d-dimensional feature space. In general though, the number of features d is quite large, making the analysis difficult. Ideally, we would like to extract the *most significant* feature (or combination of these) from the large number of original measurements available. As stated above, a typically used technique to accomplish this is PCA. In PCA, the structure of the data (i.e., location of the sample vectors $\mathbf{x}_i$ in $\mathbb{R}^d$) is preserved as much as possible in the least-square sense [15]. PCA was used by [4], [9], and [1] to facilitate the analysis and classification of LB1.

1.2 Bayes Optimal Discriminant Analysis

Discriminant analysis algorithms have their origin in the least-squares solution provided by PCA [11]. As it is well-known, PCA finds that direction within $\mathbb{R}^d$ which carries most of the covariance of the data [15]. This is readily accomplished by finding the eigenvectors of the covariance matrix Σ, i.e.,

$$\Sigma\mathbf{V} = \Lambda\mathbf{V},$$

where the columns in $\mathbf{V}$ are the eigenvectors, Λ is the diagonalized matrix of corresponding eigenvalues, and the covariance matrix is a symmetric, positive semi-definite matrix defined as

$$\Sigma = \sum_{i=1}^{n} (\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^T,$$

with n the total number of sample feature vectors, and μ their mean. Σ is, in effect, the metric we wish to maximize.

LDA can be seen as an extension of PCA where we have two metrics, $\mathbf{A}$ and $\mathbf{B}$ [28]. In general, the first metric calculates within-class variances, the second is concerned with between-class variations. The goal is to minimize the metric given by $\mathbf{A}$ while maximizing that of $\mathbf{B}$. This is equivalent to the following eigenvalue problem

$$\mathbf{A}^{-1}\mathbf{B}\mathbf{V} = \Lambda\mathbf{V}.$$

Unfortunately, this method does not work well when the two metrics disagree [21], that is, when the solution favored by the first metric $\mathbf{A}$, does not agree with that of the second metric $\mathbf{B}$. To see this, note that each of the two metrics will favor a solution (similar to PCA). That is, the metric given by $\mathbf{A}^{-1}$ will select a given linear combination of features as a solution, while that of $\mathbf{B}$ will be inclined to choose another solution. When these two solutions are similar, a compromise can be reached, and our final solution is generally appropriate. However, when the two solutions are different, LDA does not know how to resolve this issue, and will end up selecting one of the two solutions as its outcome. It has been shown by [21] that this selection is not guaranteed to be correct.

To resolve this problem, one needs to define an optimization process that does not depend on the norm of the metrics. This can be resolved by searching for that ordering of class means in the reduced space $\mathbf{v}$ where the classification error is minimized. To achieve this, the algorithm first whitens the data (i.e., transforms the original feature space to one where the data covariance

						Indonesian H. Erectus			Global pooled-sex H. sapiens		
Measurement	LB1	STS5	KNM ER 1813	D2280	KNM ER 3733	n	X	s.d.	n	X	s.d
Endocranial volume (cc) (M38)	380	485	510	775	848	15	1023.7	124.5	376	1296	149.1
Max. cranial length (M1)	143	146	149	177	182	14	195.5	11.6	3946	180.1	8.5
Max. cranial breadth (W12)	113	111	113	136	142	17	148	6.5	3937	136.3	7.2
Bi-asterion (M12)	97	90	90	101	124	15	124.7	6.6	3905	108.1	5.7
Cranial height (M17)	89	101	90	105	110	10	116.8	9	3899	132.2	7.1
Porion-vertex (M21)	75	76	77	91	83	13	102.4	7.2	1781	127	6.5
Min. frontal breadth (M9)	67	65	68	74	93	2	103	4.2	688	92.8	8.3
Orbit height (M52)	31	32	30	31	36	1	36	0	3248	33.6	2.2
Orbit breadth (M51)	32	34	33	35	42	1	43	0	3224	39.8	2.2
Nasal aperture breadth (M54)	21	26	23	27	36	1	30	0	3224	26.2	2.3
Breadth/Length*100	79	76	75.8	77	78	14	76.2	3.5	3921	75.8	4.8
Bi-auriculare/Length*100	74.1	73.2	75.8	75	73	12	71.7	3.1	3820	67.5	4.2
Height/Length*100	62.2	69.1	60.4	59	59	9	60	2.5	3885	73.5	3.8
Height/max breadth*100	78.7	90.9	79.6	77	77	10	78.1	6.1	3877	97.1	6.1
Height/Bi-auriculare *100	83.9	94.4	79.6	80	83	8	82.3	4.7	3776	109	6.4
L+B+H/3	115	119	117	135	139	9	155.1	7.4	3865	149.5	5.6

Table 1: Cranial and mandibular measurements.

matrix is equal to the identity matrix), and computes the whitened means, $\hat{\mu}_i$. This is equivalent to minimizing the within-class covariance [11]. Then, the algorithm needs to search for the solution $\mathbf{v}$ where the Bayes (classification) error is minimized. This is a convex problem with a global minimum [13]. Hence, the optimal solution can be attained without the need to rely on the norms of the two metrics defined above.

In the following, we use this Bayes optimal algorithm to determine those features that most discriminate between species and genera and conclude that, according to the data available, LB1 is most probably a descendent of an earlier than originally suspected *Homo* that developed a highly functional brain. None of the data available supports the view that LB1 corresponds to a pathologic (e.g., microcephalic) H. sapiens.

2 Who is LB1?

We divide our analysis into three studies. In the first study, we derive a morphological model of shape changes between four groups: *H. sapiens*, *H. erectus*, *early Homo* (i.e., early *H. erectus*, *H. habilis*, D2280), and *A. africanus*. This includes shape and size changes that have evolved over time, i.e., how each feature morphed over time toward that of a subsequent species. To ensure that the results of the first study are not due to a size difference, in our second experiment, we study the classification of LB1 using a set of size-normalized feature vectors. Our third and final study relates to the endocast of LB1 and those of other *Homo*. Here we look at the hypothesis that LB1 could be a diseased (pathological) *H. sapiens*.

2.1 Morphological model

To construct our morphological model, we first need to determine which features are responsible for the shape and size variations. Here, we use our Bayes optimal discriminant algorithm [13] to specify which features best describe such evolutionary path through species.

To construct this model, we used a large set of cranial and mandibular measurements presented in [4] and corresponding to the following classes: *H. sapiens*, *H. erectus*, *early Homo*, and *A. africanus*. These are summarized in Table 1. In our analysis, the mean and covariances (i.e., the first and central moments of the data) of each group are used to find the optimal basis $\mathbf{v}_1$. This dimension thus specifies that linear combination of features where samples from a common

class are close to one another and samples of different classes are as far apart from each other as possible. In this space, new samples of one class project close to their class mean. Therefore, projecting the sample vector of LB1 (which was not used to compute the discriminant space) should reveal its classification. This is shown in Fig. 1(a). According to these results, LB1 is most closely related to *early Homo* and falls nicely between these and *A. africanus* (specimen STS 5). This result should be interpreted as a first (crude) morphological model defining the variation of shape from *Australopiths* (on the right) to modern humans (on the left). This is similar to the basis vectors defining the thin-plate splines given in [3]. In our model, morphological changes are thus driven by those features identified by the discriminant vector $\mathbf{v}_1$. In particular, the weights specified by each of the dimensions in this discriminant vectors correspond to the relevance of each shape component. The most notable ones are the nasal breadth (with a weight of .2274), the orbit height (.1236) and breadth (.0931), and the porion vertex (.0918). These are, then, the evolutionary changes undergone by the specimens over time.

We compared this result to the outcome obtained with PCA, Fig. 1(b). The PCA result is indeed very different to that provided by the discriminant algorithm employed in the preceding paragraph. We see that PCA results in an unlikely classification, with LB1 being either related to *Australopiths* or as an ancestor of these (a view not supported by the other evidence). This is, hence, our first indication of how inappropriate the use of PCA may be in these applications.

To successfully determine the genus of LB1, we would like to know how similar LB1 is to STS 5 and to *early Homo*. Is LB1 closer to *Australopiths* or to the genus *Homo*? To resolve this issue, we will now include a second most discriminant basis within our morphological model. To identify this second basis $\mathbf{v}_2$, we repeat the same process described above but in the null space of $\mathbf{v}_1$. The optimal solution in the null space of $\mathbf{v}_1$ will hence be our solution $\mathbf{v}_2$, and the combination of these two discriminant features ($\mathbf{v}_1$ and $\mathbf{v}_2$) will provide the 2D model we need. This result is now shown in Fig. 1(c). In this case, this second basis vector is most marked by shape changes on the ratio between height and length (.2193), the orbit height (.1765), and the nasal breadth (.1427).

Fig. 1(c) shows the best cubic spline fit to the mean of each class. This finally defines the morphological (shape) variations from *A. africanus* to *H. sapiens* we were looking for, i.e., *our model*. As above, this morphological model describes the evolutionary shape deformations undergone over time by the considered species. In this case, however, shape variations are determined by the two bases, $\mathbf{v}_1$ and $\mathbf{v}_2$ [3, 26].

The projection of the feature vector containing the measurements of LB1 onto our 2D model is also shown in Fig. 1(c). We see that LB1 clearly fits within the morphological model describing shape changes from *A. africanus* to *early Homo*, but is clearly closer to *early Homo*. These morphological changes support the hypothesis that LB1 is related to *early Homo*.

The 2D PCA result is given in Fig. 1(d) for comparison. Again, we see that this solution results in a very unlikely classification, since (according to PCA) LB1 would be classified as an *Australopith*.

2.2 Discriminant features

A concern of the morphological model defined in the preceding section is that LB1's shape changes may actually be due to a size rather than shape difference. Size changes may be attributed to a late single mutation [6, 24], creating the illusion of moving backwards in the evolutionary model. Microcephaly is one such mutation, and several authors have claimed this was LB1's fate

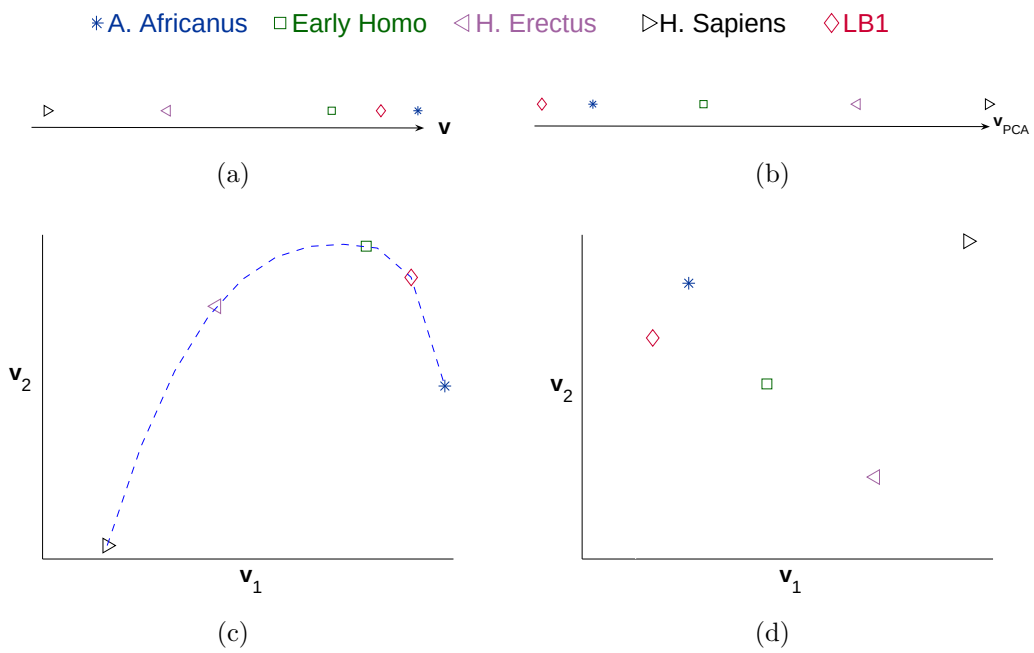
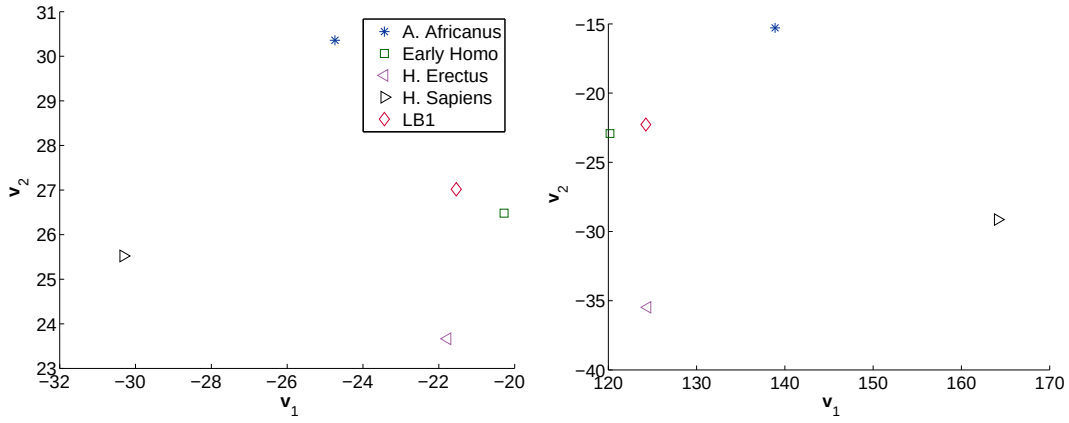


Figure 1: Morphological model. (a) The means of each class and the feature vector of LB1 (not used to generate the space) are projected onto this feature space, $\mathbf{v}$, obtained using our Bayes optimal discriminant algorithm. LB1 falls between the classes of *A. africanus* and *early Homo*. (b) For comparison the result obtained using the same procedure but with the PCA algorithm is shown here, $\mathbf{v}_{PCA}$. It is clear from this figure that PCA does not carry discriminant information, since it orders the specimens in a very unlikely sequence: from LB1 to *early Homo* with STS 5 right in between. (c) Two-dimensional representation constructed using the two most discriminant vectors obtained. Our model is obtained by fitting a cubic spline to each of the class means. This model corresponds to the morphological transitions between the shapes of different species. The model suggests LB1 is a descendant of an ancestor of *early Homo*. (d) Two-dimensional results obtained with PCA. When compared to the discriminant space shown in (c), PCA does not provide a smooth transition between taxa and presents an odd relation between *Australopithecus*, *H. sapiens* and LB1.



(a) (b)

Figure 2: (a) Shown here is the two-dimensional feature representation provided by the Bayes optimal discriminant analysis algorithm on the shape data appropriately normalized by size. LB1 is clearly classified as a relative of the *early Homo* group. We also note that the first (most) discriminant feature ($\mathbf{v}_1$) relates LB1 with *H. erectus*. (b) In this particular case, PCA provides a reasonable result. Closer analysis, however, shows that the first dimension ($\mathbf{v}_1$) in PCA attributes LB1 to *H. erectus* rather than half way from this and *early Homo*. In the 2D representation, LB1 is closer to *A. africanus* than to *H. erectus* (contrary to the results obtained with our discriminant approach).

[27, 18, 19, 14].

To test this alternate hypothesis, we build a classifier using the same data measurements used above. However, this time, we first normalized all the data with respect to height and breadth. This makes the results invariant to size and dependent on local shape only. In particular, we used the following 11 features ($d = 11$): M12/M1, M21/M1, M9/M1, M51/M1, M52/M1, M54/M1, Breath/Length, Bi-auriculare/Length, Height/Length, Height/max breadth, and Height/Bi-auriculare.

Next, we use our Bayes optimal discriminant analysis algorithm [13] to define a classifier that can assign class probabilities to each of the four main classes considered above: *H. sapiens*, *H. erectus*, *early Homo* and *A. africanus*. This classifier is constructed as follows. First we use all the normalized data but that of LB1 to find that two-dimensional representation where each class is best separated. Then, we project the sample vector representing LB1 onto this two-dimensional discriminant space. In this discriminative space LB1 is most similar to *early Homo* and shares many of the most discriminant features of *H. erectus*, Fig. 2(a). This suggest LB1 is related *early Homo* and (to a lesser degree) to *H. erectus*, consistent with the results of our previous studies.

A comparison with to results obtained with PCA, Fig. 2(b) reveals, once more, the need of the discriminant analysis defined in this paper.

2.3 Endocasts

The classification of LB1 given thus far is still eclipsed by the stone tools found at Liang Bua [22, 5]. It has been suggested that these tools are too advanced for having been developed by a

	Length	Breadth	Height	Frontal Breadth	Breadth /Length	Height /Length	Frontal Breadth /Length	(Breadth -Frontal Breadth) /Length	Height /Breadth
H.sapiens (n=8)									
MEAN	162.06	122.03	116.98	105.41	0.75	0.72	0.65	0.1	1.11
S.D.	7.07	7.32	9.4	5.48	0.03	0.04	0.02	0.03	0.06
KNM-WT	113.4	92.9	72.5	78.1	0.82	0.64	0.69	0.13	0.78
STS5	119.1	93.5	86.3	85.6	0.79	0.72	0.72	0.07	0.92
ZKD III	158.6	124.5	99.7	91.4	0.78	0.63	0.58	0.21	0.8
ZKD X	174.6	130.4	114.9	106.7	0.75	0.66	0.61	0.14	0.88
ZKD XI	165.9	127.2	103.7	97.1	0.77	0.63	0.59	0.18	0.82
ZKD XII	167.4	128	108.5	97.8	0.76	0.65	0.58	0.18	0.85
Trinil 2	156.7	126.9	95	92.5	0.81	0.61	0.59	0.22	0.75
Pygmy	165.7	123.9	116.9	102.6	0.75	0.71	0.62	0.13	0.94
LB1	119.6	102.8	81.4	77.7	0.86	0.68	0.65	0.21	0.79

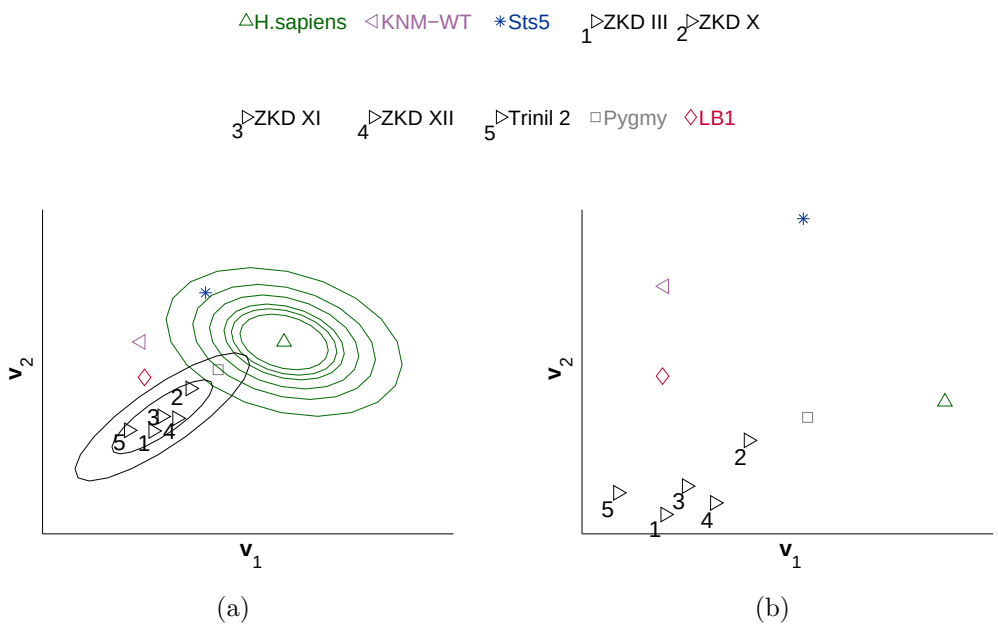
Table 2: Endocast measurements (in mm). In our test, only the ratios were considered, since these are appropriately normalized by size.

small brain *Homo* living in Flores. [5] provide some evidence to invalidate the view that these are too complex. Furthermore, [9] have suggested that LB1's brain is similar to that of *H. erectus*, with a developed Brodmann's area 10, associated to high cognitive skills. Hence, it is possible that LB1 evolved independently from other *early Homo* to a species with similar cognitive capabilities to those of *H. erectus* (albeit with a much smaller brain), or that the tools are not too complex for a small brain *Homo* to have developed them.

As already noted though, several authors [27, 18, 19, 14, 24] have argued that LB1's brain seems advanced because it actually corresponds to a *H. sapiens* that has undergone microcephaly. Although our results above seem to negate this assertion, it is still possible that LB1 was a child with morphological abnormalities and an uncanny bone structure. If this hypothesis were true, the shape of the brain endocast of LB1 should be within a relatively small deformation from that of healthy *H. sapiens*. [14] cite the common trait that the shape of microcephalics is within 5 (and possibly even 6) standard deviations (s.d.) from that of *H. sapiens*. General deformations caused by a single gene may indeed account for such variations [24].

We compiled the datasets of [9] and [8], Table 2, to test this hypothesis. Our analysis compares a set of endocasts ratio measures obtained from the following classes: *H. sapiens*, *H. erectus*, KNM WT 17000, STS 5, and LB1. In this experiment, all samples but that of LB1 were used to generate the two-dimensional discriminant space. The feature vector of LB1 was then projected onto this two-dimensional space, Fig. 3(a).

The key question to address here is whether LB1 fits within a reasonable morphological variation from *H. sapiens*. This is illustrated by the ellipsoids in Fig. 3(a). Each ellipsoid represents a percentage of the deviation from *H. sapiens* (assuming the data is Gaussian distributed). LB1 is farther than 99.8% of the variance. It is generally argued that such large changes are normally due to more than a single gene mutation such as that of microcephaly. Hence, this result suggest LB1 is not a microcephalic. LB1 also deviates from *H. erectus* by more than 99.8% of the data variance, suggesting LB1 may be a descendent of an ancestor of *H. erectus*. In Fig. 3(a) we have also projected the vector corresponding to a Pigmy. It is clear from our results that the Pigmy does correspond to a *H. sapiens*. Finally, once more, we compare our results with those provided with PCA, Fig. 3(b). Here, LB1 is incorrectly classified as a member of a different genus.



(a) (b)

Figure 3: (a) The ratio measurements from the endocasts *P. aethiopicus*, *A. africanus*, *H. erectus*, and *H. sapiens* are used to find the two most discriminant feature vectors. The feature vectors corresponding to LB1 and a pygmy are then projected onto this discriminant space. The distributions shown for *H. sapiens* and *H. erectus* correspond to the probability that the area within the ellipsoid contains all the members of its own species. The variances are obtained from the variations within *H. sapience* and *erectus*, respectively. The outer most ellipse in each class corresponds to a probability of 99.8% (assuming Gaussian distributed data). While the measurements of the pigmy fall within the normal range of *H. sapience*, LB1 does not. More interestingly, LB1 cannot be classified as a *H. erectus* either. This is consistent with our previous suggestion that LB1 shares an ancestor with *H. erectus*. (b) This two-dimensional PCA result incorrectly classifies LB1 as a *P. aethiopicus*.

3 Conclusions

The goal of this paper was to introduce a new tool that can be used to help validate or invalidate hypotheses about new specimens. This tool is based on the results of a recent algorithm [13] able to extract the one-dimensional subspace where two or more classes are optimally separated (as defined by Bayes).

We have used this statistical tool to study two hypotheses on the origins of LB1. The results of our statistical analysis agree with the hypothesis of an *early Homo* arriving in Asian [12, 7, 17, 1] and evolving into a highly functional species, *H. floresiensis*. Our morphological model and statistical analysis, suggest that the specimens that left Africa toward Flores were members of the *early Homo* with resemblance to or reminiscences of *Australopithecine*. Our results and model are in favor of those reported by Argue et al. [1], where it is suggested that LB1 may correspond to an archaic *Homo* or a specimen in morphological transition from *Australopithecus* to *Homo*.

Traditionally, statistical analysis and classification of specimens has been carried out using PCA, where those linear combinations of features carrying most of the data variance are selected to represent the fossil's data in a low-dimensional space from which conclusions can be drawn. In this paper, we have shown that this can in fact lead to misclassifications and misinterpretations of the data. The use of discriminant analysis tools had been tampered by their inefficiency at finding those one- or two-dimensional representations where the classes are optimally separated [20, 21]. In the present paper, we have shown that this can be resolved. We have also introduced a methodology to facilitate studies in paleontology and other related sciences.

The methodology used in this paper will prove instrumental for the analysis and classification of new specimens and the understanding of evolution. Our methodology can also be employed to determine the genera of a set of species or specimens or determine the species (sub-divisions) of a given genus and to test models (hypothesis) of human evolution. This can be done using conventional features or gene data. And, as more data from specimens in Dmasini and Flores become available, these techniques should prove fruitful in unveiling their underlying connections.

Acknowledgments

The authors are partially supported by the National Institutes of Health under grant R01 DC 005241 and by the National Science Foundation under grant IIS 0713055.

References

[1] Argue, D., Donlon, D., Groves, C. & Wright, R. "Homo floresiensis: microcephalic, pygmoid, Australopithecus, or Homo?" Journal of Human Evolution **51**, 360374, 2006.

[2] Bayer, U. "Pattern recognition problems in geology and paleontology," Springer-Verlag, 1985.

[3] Bookstein, F.L. "Principal Warps – Thin-plate splines and the decomposition of deformations," IEEE Trans. Pattern Anal. Mach. Intell. **11**, 567-585, 1989.

[4] Brown, P., Sutikna, T., Morwood, M.J. Soejono, R.P., Jatmiko, Saptomo, E.W. & Due, R.A. "A new small-bodied hominin from the Late Pleistocene of Flores, Indonesia," Nature **431**, 1055-1061, 2004.

[5] Brumm, A., Aziz, F., van den Bergh, G.D., Morwood, M.J., Moore, M.E., Kurniawan, I., Hobbs, D.R. & Fullagar, R. "Early stone technology on Flores and its implications for Homo floresiensis," Nature **441**, 624-628, 2006.

[6] Carroll, S.B. "Genetics and the making of Homo sapiens," Nature **422**, 849-857, 2003.

[7] Dennelle, R. & Roebroeks, W. "An Asian perspective on early human dispersal from Africa," Nature **438**, 1099-1104, 2005.

[8] Falk, D., Redmond, J.C., Guyer, J., Conroy, C., Recheis, W., Weber, G.W. & Seidler, H. "Early hominid brain evolution: a new look at old endocasts," J. Hum. Evol. **38**, 695-717, 2000.

[9] Falk, D., Hildebolt, C., Smith, K., Morwood, M.J., Sutikna, T., Brown, P., Jatmiko, Saptomo, E.W., Brunsden, B. & Prior, F. "The Brain of LB1, Homo floresiensis," Science **308**, 242-245, 2005.

[10] Fisher, R.A. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics **7**, 179-188, 1937.

[11] Fukunaga, K. "Introduction to Statistical Pattern Recognition," Academic Press (2nd edition), 1990.

[12] Gabunia, L., Vekua, A., Lordkipanidze, D., Swisher III, C.C., Ferring, R., Justus, A., Nioradze, M., Tvalchrelidze, M., Antón, S.C., Bosinski, G., Jöris, O., Lumley, M.-A.-de, Majsuradze, G. & Mouskhelishvili, A. "Earliest Pleistocene Hominid Cranial Remains from Dmanisi, Republic of Georgia: Taxonomy, Geological Setting, and Age," Science **288**, 1019-1025, 2000.

[13] Hamsici, O.C. & Martinez, A.M. "Bayes Optimality in Linear Discriminant Analysis," IEEE Trans. Pattern Anal. Mach. Intell., 2008.

[14] Jacob, T., Indriati, E., Soejono, R.P., Hsü, K., Frayer, D.W., Eckhardt, R.B., Kuperavage, A.J., Thorne, A. & Henneberg, M. "Pygmoid Australomelanesian Homo sapiens skeletal remains from Liang Bua, Flores: Population affinities and pathological abnormalities," Proc. Natl. Acad. Sci. U.S.A. **103**, 13421-13426, 2006.

[15] Jolliffe, I.T. "Principal Component Analysis," Springer-Verlag (Second Edition), 2002.

[16] Jurmain, R., Kilgore, L. & Trevathan, W. "Introduction to Physical Anthropology," Wadsworth Publishing (10^{th} edition), 2005.

[17] Lordkipanidze, D., Vekua, A., Ferring, R., Rightmire, G.P., Agusti, J., Kiladze, G., Mouskhelishvili, A., Nioradze, M., Ponce de Len, M.S., Tappen, M. & Zollikofer, C.P.E. "The earliest toothless hominin skull," Nature **434**, 717-718, 2005.

[18] Martin, R.D., MacLarnon, A.M., Phillips, J.L., Dussubieux, L., Williams, P.R. & Dobyns, W.B. "Comment on 'The Brain of LB1, Homo floresiensis'," Science **312**, 999b (online), 2006.

[19] Martin, R.D., MacLarnon, A.M., Phillips, J.L., & Dobyns, W.B. "Flores hominid: New species or microcephalic dwarf?" Anat. Rec. Part A **288A**, 1123-1145, 2006.

[20] Martinez, A.M. & Kak, A.C. "PCA versus LDA," IEEE Trans. Pattern Anal. Mach. Intell. **23**, 228-233., 2001

[21] Martinez, A.M. & Zhu, M." Where are linear feature extraction methods applicable?" IEEE Trans. Pattern Anal. Mach. Intell., **27**, 1934-1944, 2005.

[22] Morwood, M.J., Soejono, R.P., Roberts, R.G., Sutikna, T., Turney, C.S.M., Westaway, K.E., Rink, W.J., Zhao, J.X., van den Bergh, G.D., Due, R.A., Hobbs, D.R., Moore, M.W., Bird, M.I. & Fifield, L.K. "Archaeology and age of a new hominin from Flores in eastern Indonesia," Nature **431**, 1087-1091, 2004.

[23] Morwood, M.J., Brown, P., Jatmiko, Sutikna, T., Saptomo, E.W., Westaway, K.E., Due, R.A., Roberts, R.G., Maeda, T., Wasisto, S. & Djubiantono, T. "Further evidence for small-bodied hominins from the Late Pleistocene of Flores, Indonesia," Nature **437**, 1012-1017, 2005.

[24] Richards, G. "Genetic, physiologic and ecogeographic factors contributing to variation in Homo sapiens: Homo floresiensis reconsidered," J. Evol. Biol. **19**, 17441767, 2006.

[25] Rightmire, G.P., Lordkipanidze, D. & Vekua, A. "Anatomical descriptions, comparative studies and evolutionary significance of the hominin skulls from Dmanisi, Republic of Georgia," J. Hum. Evol. **50**, 115-141, 2006.

[26] Slice, D.E. "Modern morphometrics in physical antropology," New York: Kluwer Academic, 2005.

[27] Weber, J., Czarnetzki, A. & Pusch, C.M. "Comment on 'The Brain of LB1, Homo floresiensis'," Science **310**, 236b (online), 2005.

[28] Zhu, M. & Martínez, A.M. "Subclass discriminant analysis," IEEE Trans. Pattern Anal. Mach. Intell. **28**, 1274-1286, 2006.