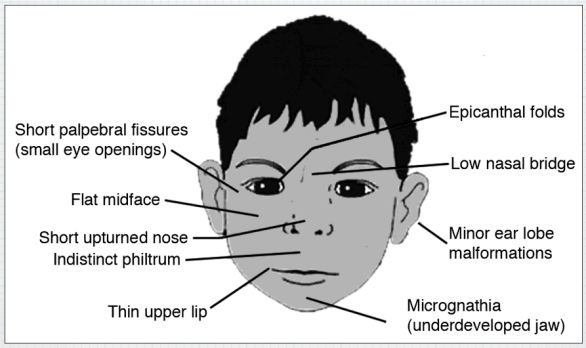
The Statistics of Shape & Recognition

John C. Russ
College of Engineering, North Carolina State Univ.

F. Brent Neal
Milliken Research Corporation

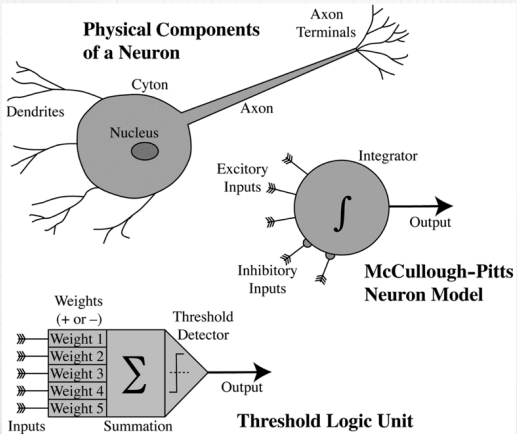
Human identification process depends on recognizing syntactical features.



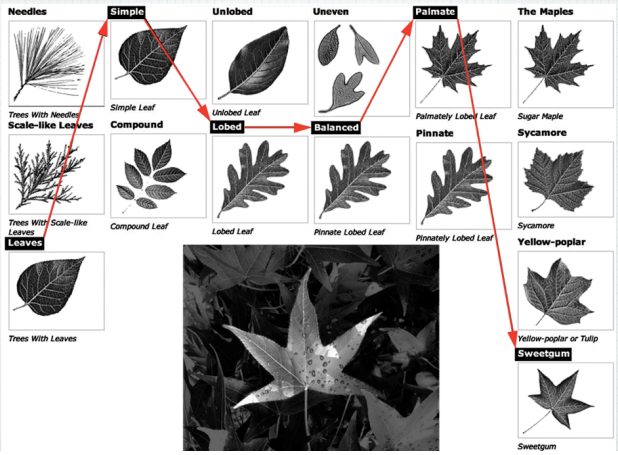
Indicators for Fetal Alcohol Syndrome

Syntactical features are "defined" by example(s) and given names for the convenience of those who have learned to recognize them. Not all of these indications will be present in each instance, and it is up to the observer to decide if enough are present, in each case, for a diagnosis.

The "Grandmother Cell" model for recognition is the basis for neural net computing.

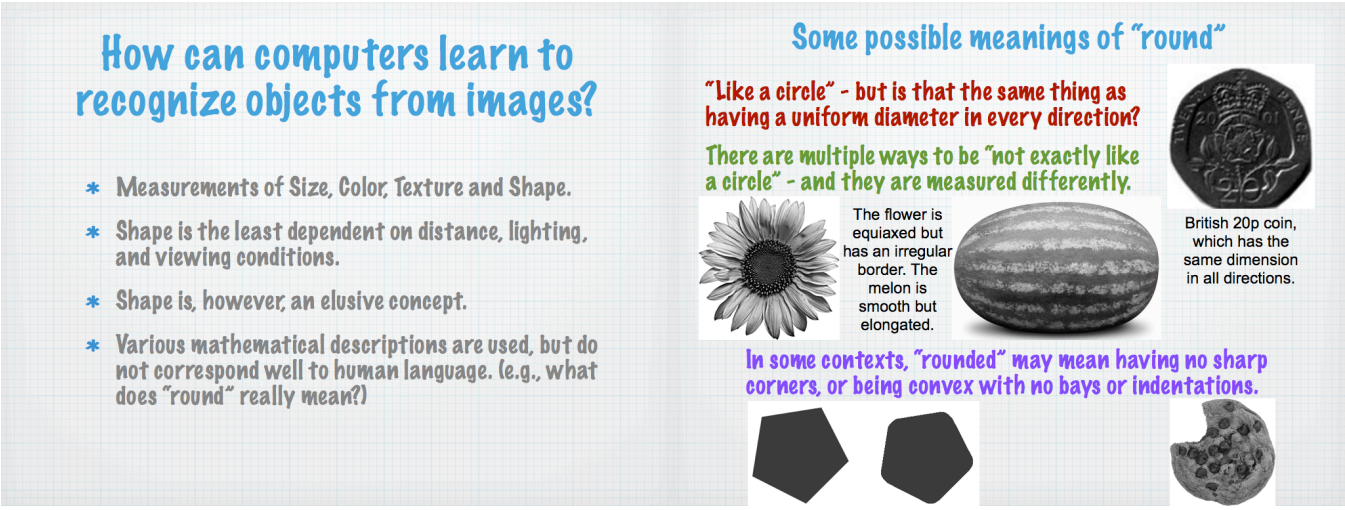


The "Field Guide" Approach



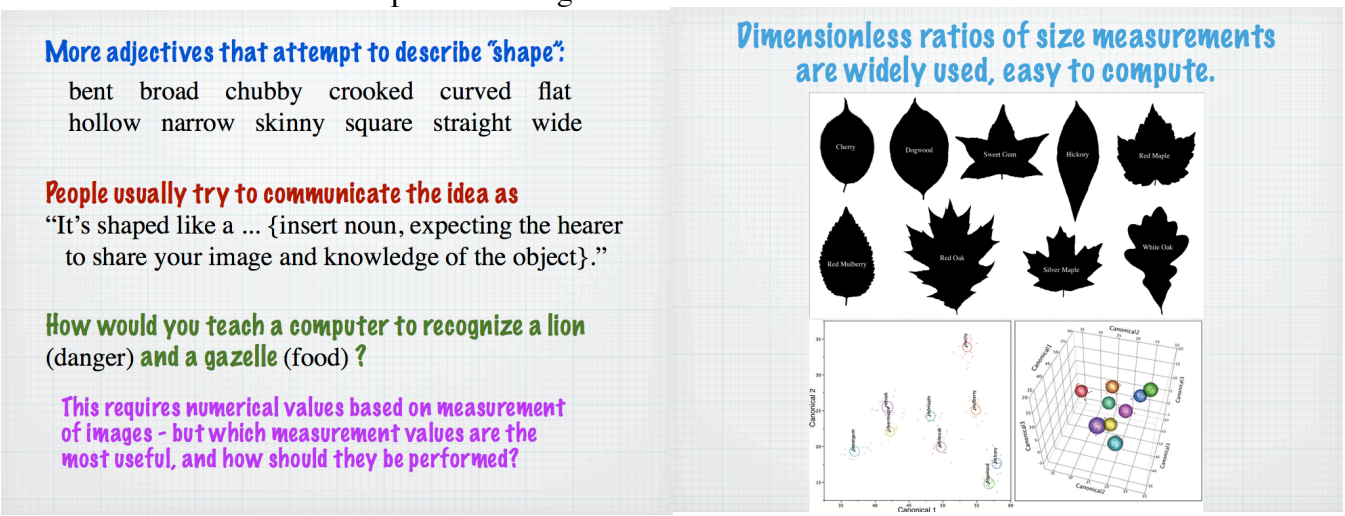
In 1942, the McCullough-Pitts model was proposed to describe recognition, and is often referred to as a "Grandmother" cell. Inputs such as "checkered apron" "short" "white hair" are positive, while "red mustache" would be negative. If the sum exceeds a threshold, you greet Grandma. The idea is readily extended to numerical inputs, which becomes the basis for neural net computing. Various methods exist for determining the weight values.

The syntactical method is also the basis for a typical field guide, e.g. to birds or flowers. In this example the pictures "define"the words. The organization and sequence of a field guide is critical to its efficient and successful use.



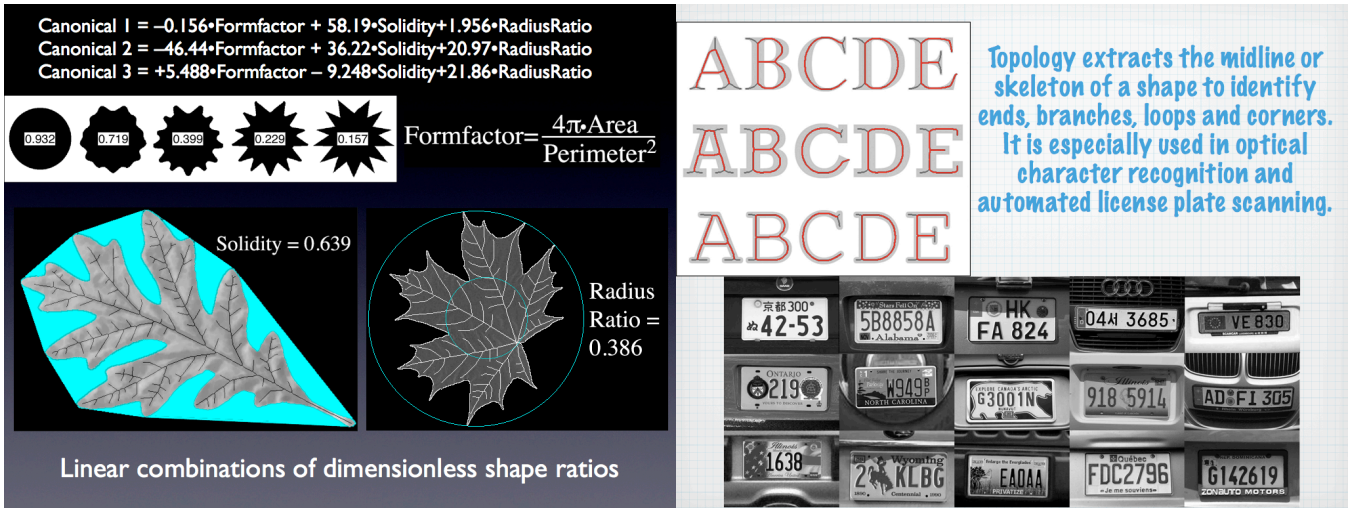
Computers can perform many measurements on digital images. Some, such as color and size, may vary. Shape is frequently the most reliable, but how is shape described and defined? Human languages do not offer many clues, and human vision is not a quantitative tool.

Consider the adjective "round" - just what does it really mean? "Like a circle." But that presumes knowledge about a circle. If something isn't EXACTLY a circle, how can the degree of difference be quantified? The problem is that there are MANY ways to be "not exactly a circle." In different contexts, the word "round" has other implied meanings.



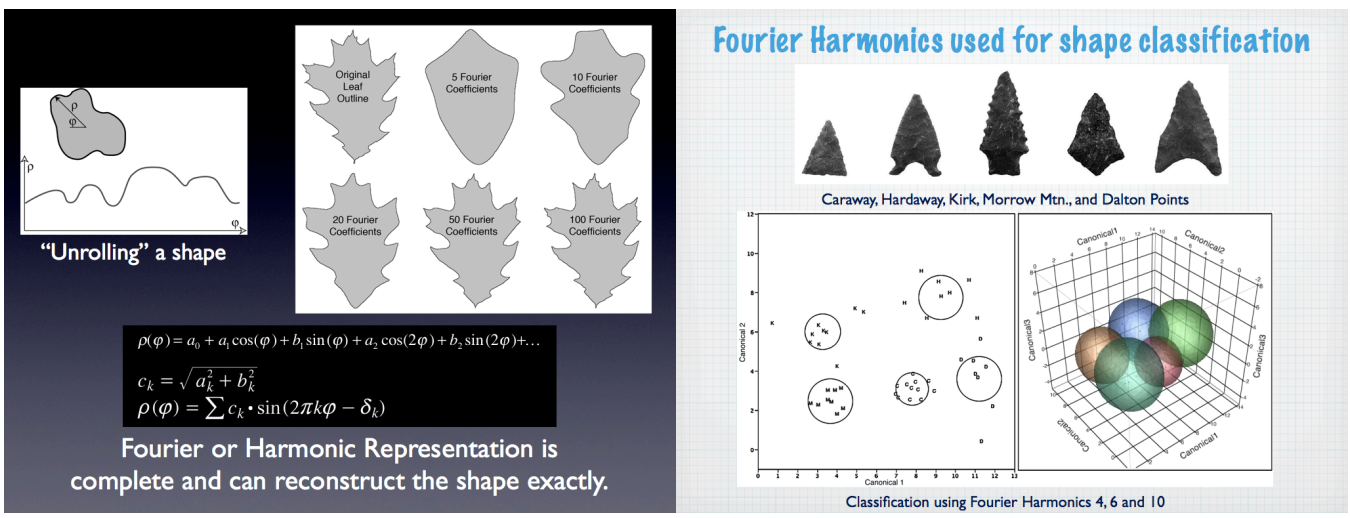
There are few adjectives that seem to describe shape, and they are all imprecise and context-dependent. So we use nouns and count on the other person to share our prior knowledge. If I say "Shaped like a volkswagon" and am thinking of the VW Bug, and you are thinking of the VW Van, we have a failure to communicate. I can show a person examples (or pictures) of objects and expect them to (eventually) be able to recognize things - that is how we teach our kids. But what about computers? Obtaining quantitative measures of shape from digital images and determining which are most useful brings us (finally) to the use of statistical analysis.

Ratios of size measurements - such as length / breadth - are formally dimensionless and very fast to compute. They are reductive, meaning that they extract just a tiny bit of the information in the original image. In this example, just a few simple measurements are sufficient to identify each leaf, based on hundreds of training examples. Ratios of distances are also used for facial recognition.



The canonical variables used in the leaf identification are simply linear combinations of various shape measurements, determined by principal components analysis and stepwise regression. The three dimensionless ratios used in the example are formfactor, solidity and radius ratio.

Topology is another shape descriptor. A and D have one loop but the A has two ends. B and 8 have two loops but the B has two corners. And so on. In spite of enormous variation in format, license plates are being read with 95% success in under a second by car-mounted cameras.



Fourier analysis "unrolls" the boundary of the shape, which repeats every 360 degrees and can be represented by a Fourier series. This series converges rapidly so that a relatively short list of numbers represents the entire shape. Statistical analysis of the terms in the series determines those that are most significant for identification.

The types of arrow points found in North Carolina are readily distinguished using just three of the Fourier coefficients, in spite of variations in materials and individual style.

Other Mathematical Shape Descriptors

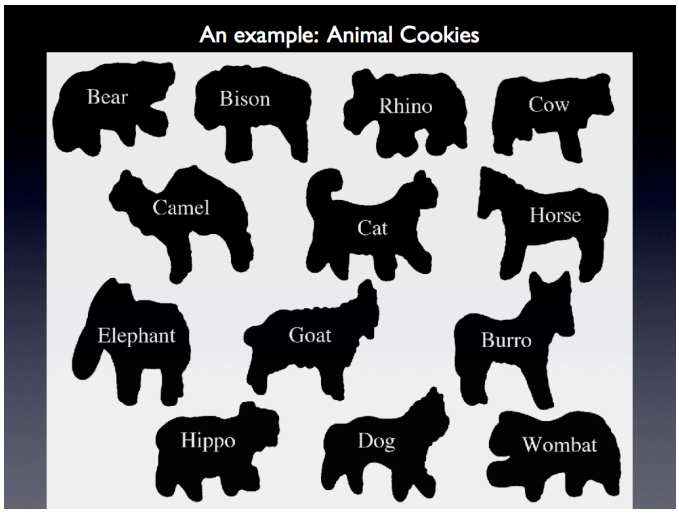
* Wavelets
* Fractal Dimension
* Curvature Scale Space
* Cross-correlation
* Moments
* etc.... see F. Brent Neal & John C. Russ, "Measuring Shape," CRC Press, 2012, isbn 978-1-4398-5598-0

Keys to success:

* An adequate training population
* Proper statistical analysis

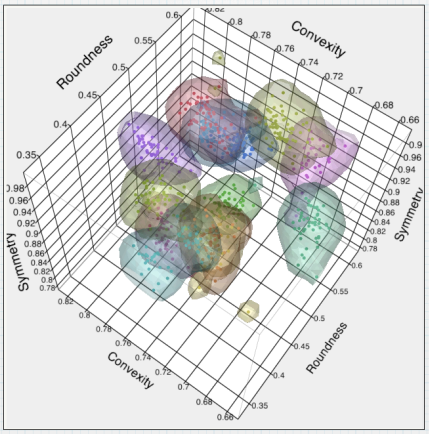
There are other mathematical shape descriptors that can also be used, and each of these has some advantages for some applications. Some are "complete" and contain enough information to recreate the original shape, some are "reductive" and extract just a few characteristics.

In any particular situation, the keys to successful shape identification and recognition are a training population that is representative, and includes extreme values that define the limits, and statistical techniques that ideally handle distributions that are rarely "normal," so that non-parametric methods are needed.



This image shows examples of 13 different classes of animal cookies. Images of hundreds of cookies from a flatbed scanner were used to measure shape. Dimensionless ratios, Fourier coefficients, and moments were all recorded for statistical analysis. (The names given to the various shapes were not supplied by the manufacturer, but instead were voted on by children.)

The following graphs and other results will be generated interactively using JMP with the files of data - dimensionless ratios, moments, and Fourier harmonic coefficients - measured on the scanned images of the cookies.



Scatterplots of the various dimensionless ratios show that the various groups are not “normal” in shape and not completely separated.

Neural

Validation: Random Holdback

Model Launch

Model NTanH(3)

Training

animal	Measures
Generalized RSquare	0.9996868
Entropy RSquare	0.990369
RMSE	0.0611745
Mean Abs Dev	0.0187474
Misclassification Rate	0.0023041
-LogLikelihood	10.622246
Sum Freq	434

Confusion Matrix

Actual / Predicted

animal	bear	bison	bobcat	camel	cat	cow	dog	donkey	elephant	goat	hippo	horse	rhino
bear	42	0	0	0	0	0	0	0	0	0	0	0	0
bison	0	32	0	0	0	0	0	0	0	0	0	0	0
bobcat	0	0	44	0	0	0	0	0	0	0	0	0	0
camel	0	0	0	29	0	0	0	0	0	0	0	0	0
cat	0	0	0	0	32	0	0	0	0	0	0	0	0
cow	0	0	0	0	0	30	0	0	0	0	0	0	0
dog	0	0	0	0	0	0	30	0	0	0	0	0	0
donkey	0	0	0	0	0	0	0	40	0	0	0	0	0
elephant	0	0	0	0	0	0	0	0	19	0	0	0	0
goat	0	0	0	0	0	0	0	0	0	33	0	0	0
hippo	0	0	0	0	0	0	0	0	0	0	34	0	0
horse	0	0	0	0	0	0	0	0	0	0	0	25	0
rhino	0	0	0	0	0	0	0	0	0	0	1	0	43

A neural-net model using all of the dimensionless shape ratios succeeds, but does little to explain which or why each one is used.

First, we will see what the dimensionless ratios can provide.

Neural

Validation: Random Holdback

Model Launch

Model NTanH(3)

Training

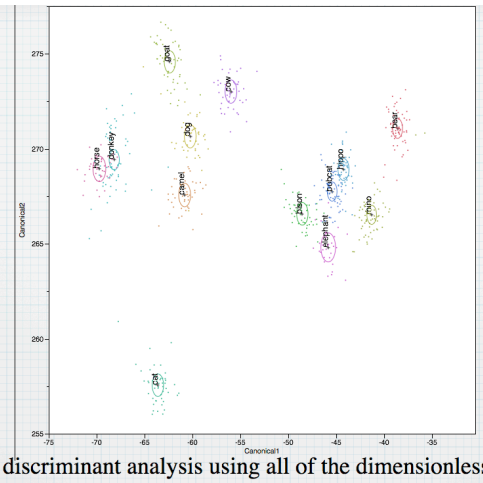
animal	Measures
Generalized RSquare	0.9994337
Entropy RSquare	0.9829155
RMSE	0.10329
Mean Abs Dev	0.0284266
Misclassification Rate	0.0092166
-LogLikelihood	18.842962
Sum Freq	434

Confusion Matrix

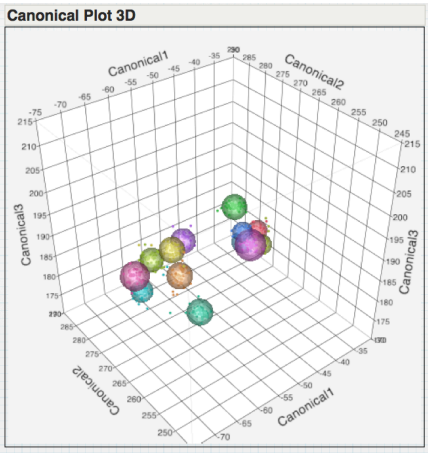
Actual / Predicted

animal	bear	bison	bobcat	camel	cat	cow	dog	donkey	elephant	goat	hippo	horse	rhino
bear	41	0	0	0	0	0	0	0	0	0	0	0	1
bison	0	32	0	0	0	0	0	0	0	0	0	0	0
bobcat	0	0	44	0	0	0	0	0	0	0	0	0	0
camel	0	0	0	29	0	0	0	0	0	0	0	0	0
cat	0	0	0	0	32	0	0	0	0	0	0	0	0
cow	0	0	0	0	0	30	0	0	0	0	0	0	0
dog	0	0	0	0	0	0	30	0	0	0	0	0	0
donkey	0	0	0	0	0	0	0	40	0	0	0	0	0
elephant	0	0	0	0	0	0	0	0	19	0	0	0	0
goat	0	0	0	0	0	0	0	0	0	33	0	0	0
hippo	0	0	0	0	0	0	0	0	0	0	34	0	0
horse	0	0	0	0	0	0	0	1	0	0	0	24	0
rhino	2	0	0	0	0	0	0	0	0	0	0	0	42

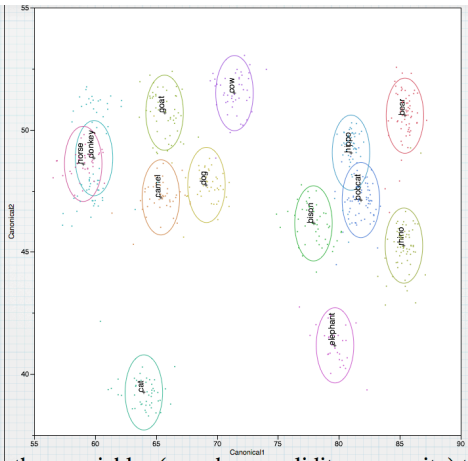
Using only four measurements (Formfactor, Roundness, Solidity, Convexity) yields equivalent success.



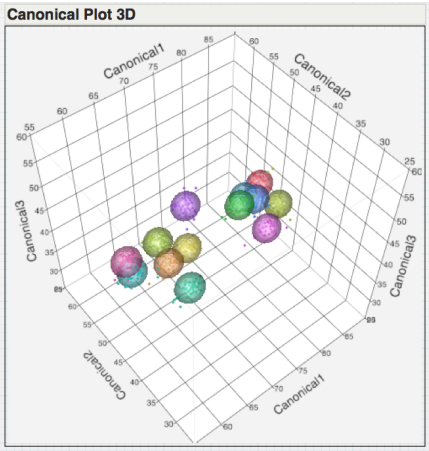
Linear discriminant analysis using all of the dimensionless ratios separates the classes with just 6 errors (all broken cookies).



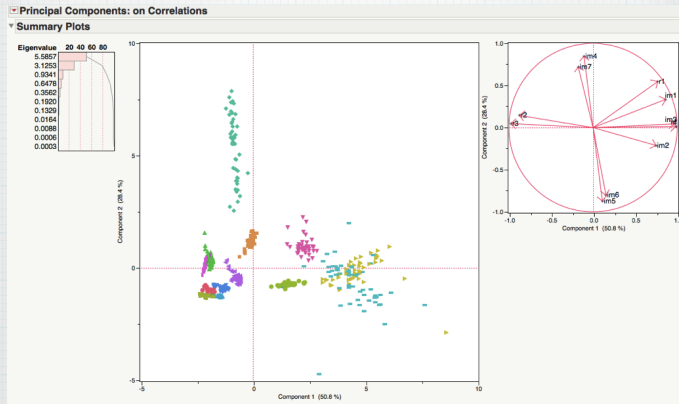
Linear discriminant analysis using all of the dimensionless ratios separates the classes with just 6 errors (all broken cookies).



With only three variables (roundness, solidity, convexity) the results are still adequate (6% outliers).

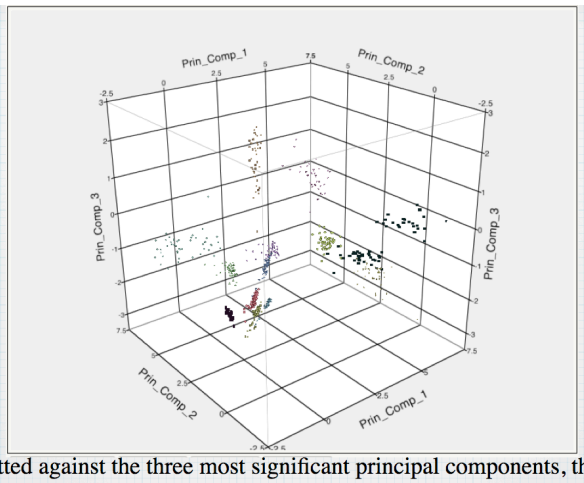


With only three variables (roundness, solidity, convexity) the results are still adequate (6% outliers).

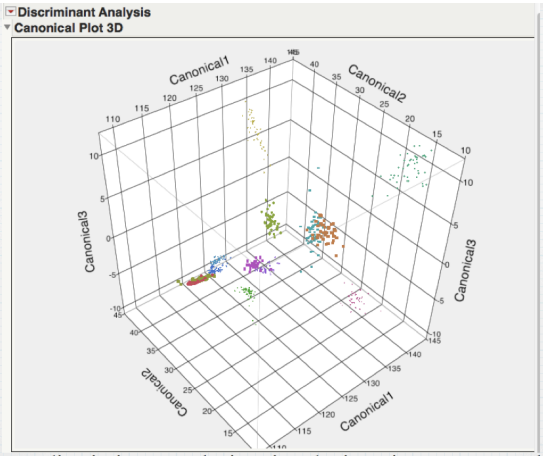


There are 11 invariant moments. Principal components analysis indicates which are most significant for distinguishing the cookies.

Next we consider the invariant moments



Plotted against the three most significant principal components, the cookie classes show good separation.



Linear discriminant analysis using the invariant moments also separates the cookie classes with less than 2% outliers.

Neural

Validation: Random Holdback

Model Launch

Model NTanH(3)

Training

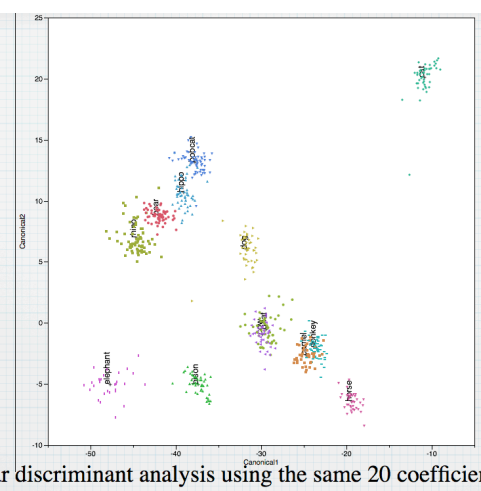
animal	Measures
Generalized RSquare	0.9997839
Entropy RSquare	0.993296
RMSE	0.0547363
Mean Abs Dev	0.0151495
Misclassification Rate	0
-LogLikelihood	7.1066147
Sum Freq	417

Confusion Matrix

Actual — Predicted

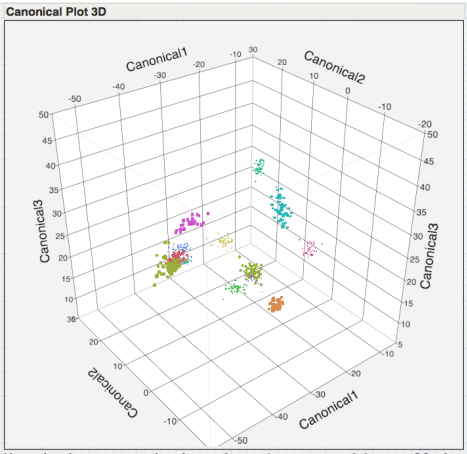
animal	bear	bison	bobcat	camel	cat	cow	dog	donkey	elephant	goat	hippo	horse	rhino
bear	40	0	0	0	0	0	0	0	0	0	0	0	0
bison	0	30	0	0	0	0	0	0	0	0	0	0	0
bobcat	0	0	41	0	0	0	0	0	0	0	0	0	0
camel	0	0	0	28	0	0	0	0	0	0	0	0	0
cat	0	0	0	0	30	0	0	0	0	0	0	0	0
cow	0	0	0	0	0	28	0	0	0	0	0	0	0
dog	0	0	0	0	0	0	28	0	0	0	0	0	0
donkey	0	0	0	0	0	0	0	38	0	0	0	0	0
elephant	0	0	0	0	0	0	0	0	19	0	0	0	0
goat	0	0	0	0	0	0	0	0	0	32	0	0	0
hippo	0	0	0	0	0	0	0	0	0	0	34	0	0
horse	0	0	0	0	0	0	0	0	0	0	0	25	0
rhino	0	0	0	0	0	0	0	0	0	0	0	0	44

A neural net model using the first 20 Fourier harmonic coefficients successfully identifies each cookie.

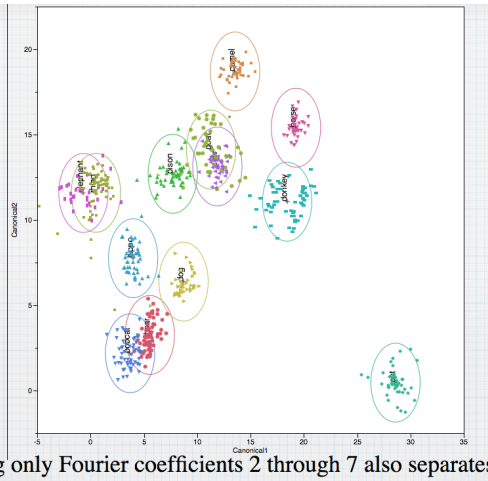


Linear discriminant analysis using the same 20 coefficients also successfully separates the cookie classes.

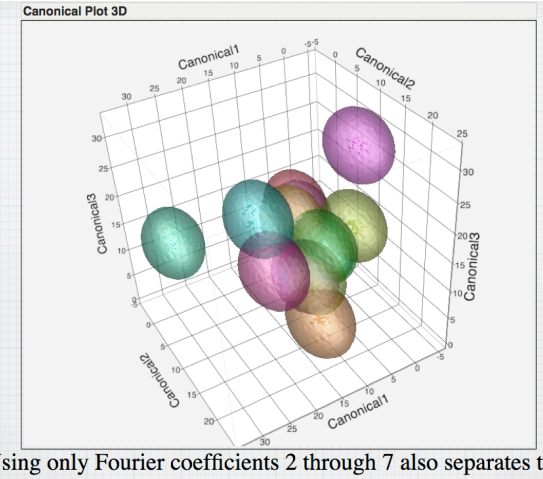
Followed by the Fourier harmonic coefficients.



Linear discriminant analysis using the same 20 coefficients also successfully separates the cookie classes.



Using only Fourier coefficients 2 through 7 also separates the classes, with less than 1% outliers.



Using only Fourier coefficients 2 through 7 also separates the classes, with less than 1% outliers.

Conclusions

- Computer object recognition is possible.
- Quantitative measurement of shape is an important element in accomplishing this.
- Using an adequate training population is essential.
- There are many applications to man-made and natural objects, including medical diagnosis.

In addition to industrial quality control, applications include diagnosing breast cancer from mammographs, identifying suspicious cells in Pap smears, and many forensic specimens such as fibers. Recognition of individual human faces is a current challenge.