Psychology of Perception

Psychology 4165, Fall 2001

Laboratory 2

Face Recognition: Are Eyes Important?

Lab 2: Signal Detection Theory and Face Recognition

Introduction

Recognition of human faces is a remarkly good skill. Even after 50 years have passed, people are able to choose which of two photographs is a high school classmate with an accuracy of almost 90 percent correct (Bahrick, Bahrick, & Wittlinger, 1974, 1975). The eyes and surrounding areas of the face are thought to be very important for face recognition (Bruce, 1988). In this lab you will test the hypothesis that obscuring the eyes makes it more difficult to later recognize a face.

There are two principle theories which describe how observers detect weak stimuli in the environment: The High Threshold Model (HTM) and Signal Detection Theory (SDT) (Krantz, 1969). In this lab you will also test the predictions of these two models for face recognition. First you will be shown a series of photographs of faces to remember. Some will have their eyes covered and some will not. Then you will be tested with another series of faces, half of which you have seen before and half of which are new. Your task is to decide after each presentation whether or not the face had been previously seen. You will respond using a six-point confidence rating scale. The use of a six-point rating scale corresponds to having five different decision criteria. From your data you will calculate five hit rate false alarm rate (HR-FAR) pairs (one pair for each of these response criteria) for the two types of faces. The two detection models will be compared to see how well each model can predict the observed data.

The objectives of this laboratory exercise are:

1. To test which model, the High Threshold Model or the Signal Detection Theory Model, better predicts your observed data in the face recognition experiment;
2. To test whether the Receiver Operating Characteristic (ROC) generated by your data conforms to that predicted by the High Threshold Model or by Signal Detection Theory;
3. To test whether or not the eyes enhance your ability to recognize faces.

Experimental Procedure

You will perform a face recognition experiment. You will first be shown a series of 64 target faces. Study each face carefully during the 2 sec exposure time. Then you will be shown a series of 128 test faces. Half of them will be target faces and half of them will be new faces that you have not previously seen. Using the six-point rating scale below rate each face on your confidence that it is a new face or a target face:

Lab 2: Signal Detection Theory and Face Recognition

1 = certain the face was not seen before
2 = perhaps the face was not seen before
3 = guessing that the face was not seen before
4 = guessing that the face was seen before
5 = perhaps the face was seen before
6 = certain the face was seen before

Don't worry about the difficulty of the task. Making these decisions sometimes is a frustrating task. Just relax and after each trial, respond with a number from 1 to 6. Be sure, though, to use all of the response categories. Record your responses in the data sheet given in Appendix I.

Data Analysis

1. Data Tabulation And Transformation

Begin your data analysis with a series of transformations of the raw data through the following sequence:

1. response scale frequencies;
2. response scale probabilities;
3. cumulative probabilities;
4. z-score transformation of the cumulative probabilities.

First: Count up the number of times you used each of the rating categories for signal trials and then count them again for noise trials. These are the response scale frequencies. Record them in the column labeled "Rating Frequency" in Appendix II.

Second: Convert the rating scale frequencies into rating scale probabilities by dividing each rating scale frequency by the total number or signal trials or the total number of noise trials as is appropriate. The probabilities should be computed to **4 significant digits**. Record these probabilities in the column labeled "Rating Probability."

Third: Convert the rating probabilities into cumulative probabilities. Start cumulating with the highest confidence signal present category (category 6) and work downward. The first cumulative probability will be the probability of saying 6. The second cumulative probability will be the probability of 6 plus the probability of 5, and so on. The cumulative probabilities for the signal trials are interpreted as hit rates (HR) and the cumulative probabilities for the noise trials are interpreted as false alarm rates (FAR). Note that the last cumulative probability should be 1.00. If it is not, check your calculations. Calculate these probabilities to **4 significant digits**.

Fourth: Transform the cumulative probabilities for scale categories 6, 5, 4, 3, and 2 (but NOT category 1) into z-scores, using the z-score tables found in all statistic books. Remember, probabilities less than 0.5 have negative z-scores and probabilities greater than 0.5 have positive z-scores. There is no z-score for probabilities of 0 or 1. If a cumulative probability from category 6, 5, 4, 3, or 2 (but NOT 1) is 0, use the z-score for 0.5/N (in this case 0.007692). If one of these cumulative probabilities is 1, use the z-score for (1–(0.5/N)) (in this case 0.9923). Arrange your data and their transformations in a table like the one shown in Appendix II.

We have provided you with a computer program called rscore+ which will compute all the data transformations for you. Double click on the data template file (dataTemplate.dat) and enter the frequency of each confidence rating for the three types of stimulus trials: new faces, old faces that had no eyes, old faces with eyes. Insert your name in the title line. Make sure that you save this file as text only. Now double click on rscore+(fat), type in your input file name. You should analyze your data twice: once for the Gaussian signal detection model and once for the high threshold model. The program will produce three output files having the same name as your input file, but with different extensions:

1. myfile.doc — a Word file with the printed output of each analysis
2. myfile.grf — a text file to be imported into Kaleidagraph for making graphs
3. myfile.alt — a results file for statistical analysis (you will not use this file)

2. ROC Analysis

The high threshold model of detection predicts that the ROC will be a straight line when plotted in probability coordinates; The signal detection theory model predicts that the ROC will be a straight line when plotted in z-score coordinates. The first step, therefore, is to plot your HR FAR pairs on two types of graphs: one with linear probability coordinates and one with z-score coordinates. Import the appropriate *.grf file into KaleidaGraph and you will have all the data you need for the various graphs. The probability graph should be square, with horizontal (FAR) and vertical (HR) axes ranging from 0 to 1. The z-score graph also should be square, with horizontal (z(FAR)) and vertical (z(HR)) axes ranging from -4.0 to +4.0 (see Appendix VII for examples). Calculate the best-fitting ROC predicted by the high threshold model (see Appendix III) and the best-fitting ROC predicted by the signal detection theory model (see Appendix IV). These results are given in your printed output. Draw these best-fitting ROC on the appropriate graphs. Which model, the high-threshold or the signal detection, gives a better description of your data? Why?

3. Sensory Process Sensitivity Analysis

Both models have a sensory process and a decision process. The sensory process is supposed to operate independently from the decision process (it is supposed to be unaffected by different decision criteria). You can test this claim of independence by calculating the value of the sensitivity index for each of the HR, FAR pairs in your data. The index p is from the high threshold model; the index d_a is from the signal detection theory model. Calculate five values of p (see Appendix III) for each of the HR, FAR pairs and plot these on a graph. This graph should have decision criterion (q = FAR) on the horizontal axis and will have p plotted on the vertical axis. Also calculate five values of d_a (see Appendix IV) for each of the z(HR), z(FAR) pairs and plot these on a graph. This graph should have decision criterion (X_C = –z(FAR)) on the horizontal axis and will have d_a plotted on the vertical axis. Which set of sensitivity indices is independent of decision criterion: those from the high threshold model or those from signal detection theory?

4. Goodness of Fit

The ultimate test of each model is by how well it predicts the observed data. Since the parameters of each model were computed using a maximum-likehood regression, the appropriate test is to calculate how well each model predicts the six response frequencies under the signal conditions. The goodness of fit of these predictions can be formally computed using the chi square (χ^2) test. Thevalue of χ^2 for each model is given on your printed output pages. If you want to do it by hand, the first step would be to compute the response frequencies predicted by each model. The procedure is to start with the ROC and work backward, just the reverse of how you transformed your original data. Start with the five observed false alarm rates (for the High Threshold Model) and the five z-score of the false alarm rates (for the signal detection model) and use the appropriate formula for the ROC to compute the predicted hit rates and z-score of the hit rates. Enter these in the tables of Appendix VIa and VIb on the right hand side. Then transform the predicted Hit Rates into the Predicted Response Probabilities and these into the Predicted Response Frequencies. For each model compute the value of χ^2 and test the statistical significance. Note that the degrees of freedom for the test of the High Threshold Model is different than the one for the Signal Detection Model. Which model fits the observed data better (has the lower χ^2)? Can you reject the hypothesis that one or both models is a good predictor of the data?

5. Group Data

When you have finished your rscore+ analysis, enter your important results on the summary sheet in Appendix VII. Transfer these results to the group sheet that Katharine will have for you so that a group data file can be prepared for further analysis. You will want to use this StatView file to test two hypotheses:

1. Seeing the eyes improves face recognition;

2. The signal detection model is better than the high threshold model

Lab Report

Your lab report should contain four parts: Introduction, Methods, Results, and Discussion. In the introduction explain why you did the experiment. In the methods section describe what you did. In the results section present your findings, including graphs of your data. Put the tables of your raw and transformed data in an appendix. Your conclusions should be based on your analysis, not your unsupported speculation. In the discussion you can let your creativity run wild. Give the reader your interpretation of the results. Discuss any implications and leads for further research. Laboratory reports must be typed, double-spaced on 8.5 x 11 paper with at least 1 inch margins. Conciseness and clarity are extremely important characteristics of good scientific writing. Strive for them. Worth 40 points. **Due in lab on 9 and 11 of October 2001**.

Lab 2: Signal Detection Theory and Face Recognition

Appendix I: Data Sheet

Trial	Rating	Target
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		
35		
36		
37		
38		
39		
40		
41		
42		
43		

Trial	Rating	Target
44		
45		
46		
47		
48		
49		
50		
51		
52		
53		
54		
55		
56		
57		
58		
59		
60		
61		
62		
63		
64		
65		
66		
67		
68		
69		
70		
71		
72		
73		
74		
75		
76		
77		
78		
79		
80		
81		
82		
83		
84		
85		
86		

Trial	Rating	Target
87		
88		
89		
90		
91		
92		
93		
94		
95		
96		
97		
98		
99		
100		
101		
102		
103		
104		
105		
106		
107		
108		
109		
110		
111		
112		
113		
114		
115		
116		
117		
118		
119		
120		
121		
122		
123		
124		
125		
126		
127		
128		

Appendix II: Sample Tables For Presentation Of Your Data

Data For Signal Trials (old faces with eyes)

Rating Category	Rating Frequency	Rating Probability	Cumulative Probability	z-Score of Probability
6				
5				
4				
3				
2				
1			1.0000	
Total	32			

Data For Signal Trials (old faces without eyes)

Rating Category	Rating Frequency	Rating Probability	Cumulative Probability	z-Score of Probability
6				
5				
4				
3				
2				
1			1.0000	
Total	32			

Data For Blank Trials (new faces)

Rating Category	Rating Frequency	Rating Probability	Cumulative Probability	z-Score of Probability
6				
5				
4				
3				
2				
1			1.0000	
Total	64			

Lab 2: Signal Detection Theory and Face Recognition

Appendix III: High-Threshold Model Formulae

The ROC predicted by the high-threshold model is a straight line passing through 1, 1 and having a y-intercept of *p*:

$$HR = p + (1 - p) \cdot FAR \qquad \text{Receiver Operating Characteristic} \qquad (2.1)$$

The sensitivity of the sensory process is indicated by *p* and the guessing rate of the decision process by *q*. These parameters may be computed from a hit rate, false alarm rate pair:

$$p = \frac{HR - FAR}{1 - FAR} \qquad \text{Sensitivity index} \qquad (2.2)$$

$$q = FAR \qquad \text{Guessing probability} \qquad (2.3)$$

To compute the least-squares estimate of p for a series of hit rate, false alarm rate pairs:

$$p = \frac{\sum Y - \left(\sum X \cdot Y\right) - \sum X + \sum X^2}{N - \left(2 \cdot \sum X\right) + \sum X^2} \qquad (2.4)$$

where $X =$ FAR and $Y =$ HR. The value of *p* from Equation 2.4, is based on the best fitting ROC line through the HR, FAR pairs, using the least-squares criterion of curve fitting.

Appendix IV: Signal Detection Theory Formulae

The ROC predicted by signal detection theory is a straight line in z-score coordinates with slope b and y-intercept a:

$$z(HR) = a + b \cdot z(FAR) \qquad \text{Receiver Operating Characteristic} \qquad (3.1)$$

The sensory process is described by the mean, μ_S, and standard deviation, σ_S, of the signal distribution. The decision process is described by X_C, the decision criterion. These parameters are computed by the following equations:

$$\mu_s = \frac{a}{b} \qquad \text{Mean of the signal distribution} \qquad (3.2)$$

$$\sigma_s = \frac{1}{b} \qquad \text{Standard deviation of signal distribution.} \qquad (3.3)$$

$$X_c = -z(FAR) \qquad \text{Decision criterion} \qquad (3.4)$$

The sensitivity or accuracy of the sensory process is expressed in terms of d_a or A_z:

$$d_a = \frac{\mu_s - \mu_n}{\sqrt{\left(\frac{\sigma_s^2 + \sigma_n^2}{2}\right)}} \qquad \text{Sensitivity index} \qquad (3.5a)$$

$$d_a = \frac{\sigma_s \cdot z(HR) - z(FAR)}{\sqrt{\left(\frac{\sigma_s^2 + \sigma_n^2}{2}\right)}} \qquad \text{Sensitivity index} \qquad (3.5b)$$

$$A_z = z^{-1}\left(\frac{d_a}{\sqrt{2}}\right) \qquad \text{Sensitivity index} \qquad (3.6)$$

where $z^{-1}(\,)$ is the inverse z-transform based on the unit normal Gaussian distribution. A_z is the area under the ROC when it is plotted in probability coordinates.

Lab 2: Signal Detection Theory and Face Recognition

Appendix Va: Predictions of the High Threshold Model and Goodness of Fit

High Threshold Model Predictions

Rating	Observed Response Frequency	Predicted Response Frequency	Predicted Response Probability	Predicted Hit Rate Probability	Observed False Alarm Rate
6					
5					
4					
3					
2					
1				1.0000	1.0000
Total	64	64			

$$\chi^2 = \sum_{rating=1}^{6} \left[\frac{(O-P)^2}{P} \right]$$

$$df = 6 - 1 - 1 = 4$$

Lewis O. Harvey, Jr. – Instructor
Katharine L. Tepe – Assistant
Room MUEN D-156, 11:00–13:50 T&Th

Lab 2: Signal Detection Theory and Face Recognition

Appendix Vb: Predictions of the Signal Detection Model and Goodness of Fit

Signal Detection Theory Predictions

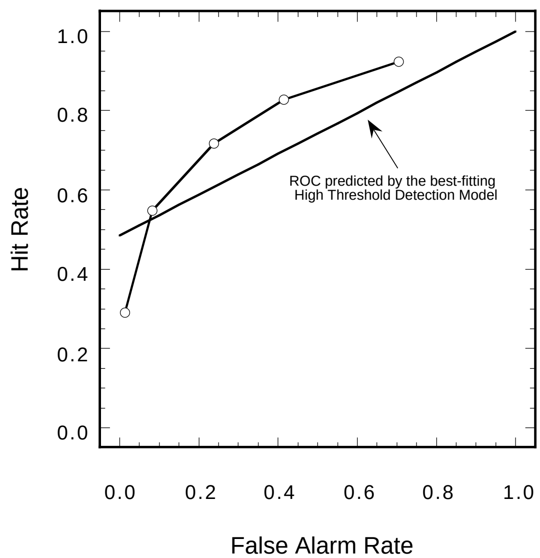
Rating	Observed Response Frequency	Predicted Response Frequency	Predicted Response Probability	Predicted Hit Rate Probability	Predicted z-score Hit Rate	Observed z-score False Alarm Rate
6						
5						
4						
3						
2						
1				1.0000	———	———
Total	64	64				

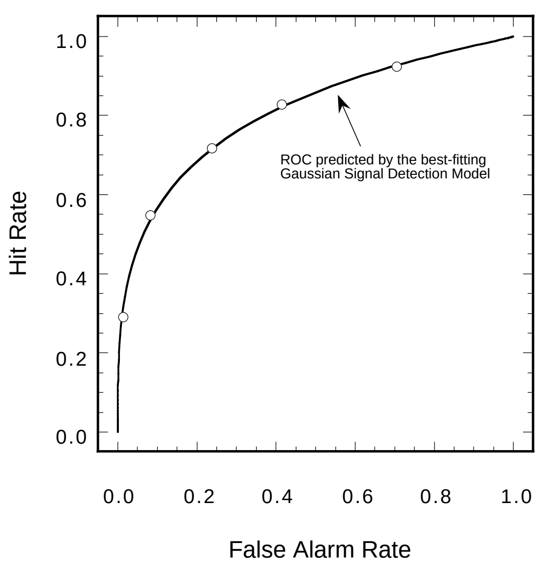
$$\chi^2 = \sum_{rating=1}^{6} \left[\frac{(O-P)^2}{P} \right]$$

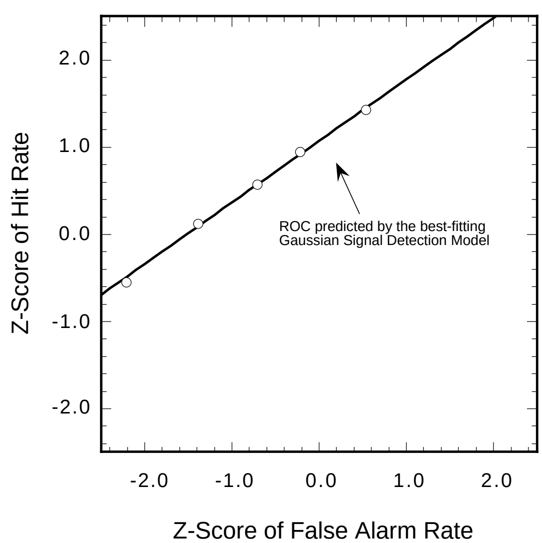
$$df = 6 - 1 - 2 = 3$$

Lab 2: Signal Detection Theory and Face Recognition

Appendix VI: Recommended Format for Graphs







Lab 2: Signal Detection Theory and Face Recognition

Appendix VII: Summary of Results

SDT d_a		SDT A_z		Goodness-of-Fit	
Without eyes	With eyes	Without eyes	With eyes	Chi-Square	probability

HTM p		Goodness-of-Fit	
Without eyes	With eyes	Chi-Square	probability

Lab 2: Signal Detection Theory and Face Recognition

References

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1974). Long-term memory: Those unforgettable high-school days. *Psychology Today, 8*(7), 50-56.

Bahrick, H. P., Bahrick, P. O., & Wittlinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*(1), 54-75.

Bruce, V. (1988). *Recognising faces*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Egan, J. P. (1975). *Signal Detection Theory and ROC Analysis*. New York: Academic Press.

Green, D. M., & Swets, J. A. (1966/1974). *Signal detection theory and psychophysics* (A reprint, with corrections of the original 1966 ed.). Huntington, New York: Robert E. Krieger Publishing Co.

Harvey, L. O., Jr. (1992). The critical operating characteristic and the evaluation of expert judgment. *Organizational Behavior & Human Decision Processes, 53*(2), 229-251.

Krantz, D. H. (1969). Threshold theories of signal detection. *Psychological Review, 76*(3), 308-324.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

McNicol, D. (1972). *A primer of signal detection theory*. London: George Allen & Unwin.

Simpson, A. J., & Fitter, M. J. (1973). What is the best index of detectability? *Psychological Bulletin, 80*(6), 481-488.

Swets, J. A. (1961). Is there a sensory threshold? *Science, 134*, 168-177.

Swets, J. A. (1986a). Form of empirical ROC's in discrimination and diagnostic tasks: Implications for theory and measurement of performance. *Psychological Bulletin, 99*(2), 181-198.

Swets, J. A. (1986b). Indices of discrimination or diagnostic accuracy: Their ROC's and implied models. *Psychological Bulletin, 99*(1), 100-117.

Swets, J. A., Tanner, W. P., Jr., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review, 68*(5), 301-340.

Lab 2: Signal Detection Theory and Face Recognition

(Egan, 1975; Green & Swets, 1966/1974; Harvey, 1992; Krantz, 1969)

(Macmillan & Creelman, 1991; McNicol, 1972; Simpson & Fitter, 1973)

(Swets, 1961, 1986a, 1986b; Swets, Tanner, & Birdsall, 1961)