

Appendix A

We are interested in the probability of classification over ten categories, but our subject data was gathered for a 2AFC task. Here we demonstrate how a simple conversion can be used to construct a full confusion matrix for the ten scene categories.

In our 2AFC task, scene A is shown with word choices A and K, where K is one of the other 9 scene categories. The proportion of times A is correctly chosen as the scene shown is equal to

$$\frac{Hits}{Trials} = \frac{P_A}{P_A + P_K}$$

and the proportion of times K is incorrectly chosen is equal to

$$\frac{Misses}{Trials} = \frac{P_K}{P_A + P_K}$$

Let N_K be the ratio of misses to hits

$$N_K = \frac{Misses}{Hits}$$

$$\Rightarrow P_K = N_K P_A$$

Given multiple scene choices, the sum of probabilities of choosing any scene must sum to one. In other words, the entries along one row in the final confusion matrix must sum to one.

$$P_A + \sum_{K \neq A} P_K = 1$$

Given this constraint and the ratio of hits to misses for a type of trial, we can convert any 2AFC observance into a row in our confusion matrix.

$$P_A + \sum_{K \neq A} N_K P_A = 1$$

$$\Rightarrow P_A = \frac{1}{\left[1 + \sum_{K \neq A} N_K\right]}$$

To better illustrate the conversion, here is an example. Assume we only have three scene categories, A, B and C, and that we obtain the results in Table A.1 with a 2AFC task. Let's compute the row in the confusion matrix for scene A. In other words, given three scene choices, how often is scene A classified as A, B and C?

$$N_B = \frac{20.0}{80.0} = 0.250 \text{ and } N_C = \frac{5.0}{95.0} = 0.053$$

$$P_A = \frac{1}{\left[1 + (N_B + N_C)\right]} = \frac{1}{1 + 0.250 + 0.053} = 0.767$$

$$P_B = N_B P_A = (0.250)(0.767) = 0.191$$

$$P_C = N_C P_A = (0.053)(0.767) = 0.041$$

When scene A is shown, it will be classified correctly as A 77% of the time, incorrectly as B 19% of the time and incorrectly as C 4% of the time. The final confusion matrix for this example can be found in Table A.2.

Table A.1. Sample data for a hypothetical 2AFC scene matching task. The three scene classes are A, B, and C. They would each be shown with the matching word choice and the word choice in the second column. The third column is the percentage of trials in which the subjects choose the correct word choice.

Scene shown	Other Word Choice	% Correct Measured
A	B	80.0
A	C	95.0
B	A	85.0
B	C	90.0
C	A	75.0
C	B	98.0

Table A.2. Confusion matrix for three scene classes, converted from the sample 2AFC data in Table A.1.

Scene shown	Classified as A	Classified as B	Classified as C
A	76.7	19.1	4.1
B	13.7	77.7	8.6
C	24.6	1.5	73.9

Appendix B

Table B.1 shows the full confusion matrices for subject and texture model data for basic-level categorization of the scenes. Across each row is the percentage of times an image of a given class is categorized as each of the 10 categories. Correct performance measures occur along the diagonal. The Monte Carlo method was used to compute the standard deviations and mean observations. Simulations were run in which half of the data was used to compute the mean of an observation. 100 observations were used. Table B.2 shows the confusion matrices for subject and texture model data for superordinate-level categorization of the scenes. These matrices are computed by summing across the basic-level categories.

Table B.1. Basic-level confusion matrices estimated from 2AFC data for subjects and the texture model.

Scene Shown	Category Selected (%)									
	Beach	Forest	Mtn	City	Farm	Street	Bath	Bedrm	Kitchn	Livingrm
<i>Subject Classifications at 37ms</i>										
Beach	18.7±2.8	8.2±4.6	10.7±6.5	3.7±2.6	10.7±4.9	6.5±3.4	10.3±5.2	14.5±5.5	8.3±3.9	8.4±3.9
Forest	7.2±4.3	21.0±3.6	30.6±10.6	5.1±3.1	11.1±5.2	7.0±3.8	7.3±4.3	1.1±1.3	3.8±2.3	5.8±3.7
Mtn	4.9±3.6	10.6±5.3	33.2±4.7	8.5±4.9	3.5±2.7	5.8±3.6	16.1±7.1	6.2±4.0	5.7±4.2	5.6±3.3
City	10.0±5.2	10.9±5.4	5.8±2.7	22.5±3.6	14.0±6.4	6.0±3.5	5.1±3.2	16.5±8.3	3.8±2.6	5.4±3.1
Farm	11.8±5.4	16.8±7.2	9.4±4.8	7.4±4.4	22.6±3.5	7.1±3.5	5.1±3.3	7.6±3.7	6.8±3.3	5.4±3.4
Street	7.9±3.7	14.7±6.3	6.7±3.2	11.7±5.1	11.1±4.9	23.1±3.5	3.7±2.6	2.6±1.9	8.2±4.7	10.1±4.7
Bath	3.0±2.0	2.8±1.9	1.8±1.6	6.4±3.5	0.9±1.0	0.9±1.1	18.9±4.0	19.3±8.8	26.0±12.0	20.0±8.8
Bedrm	4.6±2.9	3.5±2.5	5.6±3.1	4.8±3.2	4.4±2.7	5.2±3.1	1.4±1.5	28.3±6.4	8.7±4.5	33.5±10.6
Kitchn	4.5±3.0	1.2±1.2	4.3±2.5	7.6±4.7	5.1±2.9	7.1±4.0	11.2±5.5	18.9±9.2	21.1±3.5	19.0±8.0
Livingrm	6.1±3.5	6.6±3.9	2.2±2.3	9.0±5.2	6.8±4.1	18.3±7.9	4.0±3.0	5.8±3.4	3.9±3.4	37.3±5.5
<i>Subject Classifications at 50ms</i>										
Beach	18.7±2.8	16.1±5.5	9.9±3.9	17.7±6.7	11.3±4.9	5.1±2.2	2.1±1.3	3.2±2.0	2.5±1.5	13.5±5.4
Forest	18.6±6.9	32.2±4.8	6.7±3.5	5.6±2.9	17.7±7.7	5.8±3.5	5.2±2.8	2.7±1.7	4.2±2.7	1.3±1.3
Mtn	1.7±1.5	15.9±6.3	38.7±4.6	13.1±5.0	9.8±4.5	4.6±2.9	6.2±3.4	3.1±2.0	1.7±1.6	4.9±2.9
City	5.0±3.2	6.7±3.8	9.9±5.1	41.9±4.8	7.0±3.7	5.8±3.9	5.4±3.4	5.5±3.1	7.0±3.9	5.8±3.0
Farm	16.7±5.7	9.9±4.8	16.4±6.1	4.8±2.7	32.1±4.2	5.9±3.2	3.9±2.4	5.5±3.6	2.6±2.1	2.2±1.8
Street	2.2±1.7	3.5±2.5	6.1±3.2	29.9±10.0	13.6±4.9	28.4±4.4	3.3±2.2	6.8±3.8	2.5±2.1	3.8±2.5
Bath	4.9±2.2	1.0±1.2	2.2±1.8	5.7±2.9	2.1±1.5	5.6±3.2	26.8±3.5	13.4±5.2	24.1±8.2	14.3±5.5
Bedrm	0.0±0.0	0.0±0.0	4.2±2.8	4.1±2.5	7.9±4.1	7.7±3.9	4.4±2.7	33.6±5.2	10.5±4.0	27.6±9.1
Kitchn	3.4±2.2	3.5±2.1	2.6±1.8	1.2±1.1	7.2±3.2	7.8±3.3	9.5±3.6	11.0±4.5	27.8±4.0	26.2±9.0
Livingrm	3.8±2.6	6.2±3.8	3.7±2.6	10.3±4.3	6.5±3.6	4.3±3.2	1.7±2.2	4.0±2.7	9.6±4.5	49.9±6.3
<i>Subject Classifications at 62ms</i>										
Beach	41.6±8.5	3.1±3.3	27.2±11.5	7.0±5.7	7.7±5.3	0.0±0.0	2.7±3.5	3.9±4.1	3.2±3.8	3.5±3.8
Forest	4.6±5.1	59.8±10.5	15.2±10.3	10.4±8.7	0.0±0.0	0.0±0.0	4.9±5.6	0.0±0.0	5.2±5.5	0.0±0.0
Mtn	5.1±5.5	5.6±5.5	65.1±9.6	5.7±5.3	0.0±0.0	0.0±0.0	4.3±4.4	4.3±5.3	10.0±8.5	0.0±0.0
City	3.4±3.8	8.0±6.8	8.3±6.9	49.6±9.9	14.7±9.0	0.0±0.0	3.7±3.7	8.8±7.4	3.6±3.9	0.0±0.0
Farm	9.7±6.3	2.2±3.1	15.1±9.0	10.2±6.7	36.7±6.6	11.3±7.6	5.3±4.9	0.0±0.0	2.7±3.2	6.8±6.1
Street	8.0±7.0	0.0±0.0	4.1±4.1	17.5±9.7	12.5±7.8	46.4±8.5	4.1±4.5	3.9±4.1	3.5±4.2	0.0±0.0
Bath	0.0±0.0	2.8±3.4	3.2±3.4	8.6±6.9	3.9±3.7	3.6±3.3	43.1±6.9	8.5±5.9	22.8±10.0	3.4±3.8
Bedrm	0.0±0.0	9.2±6.0	6.0±4.6	10.4±7.1	13.7±6.9	3.1±3.3	7.3±5.4	36.6±6.8	6.0±5.1	7.9±10.5
Kitchn	5.0±4.2	0.0±0.0	19.2±10.1	2.3±2.9	8.4±5.9	2.1±2.7	7.8±5.5	10.0±7.1	32.3±6.7	12.9±7.1
Livingrm	4.0±4.8	0.0±0.0	4.0±4.1	0.0±0.0	3.8±3.9	5.0±4.8	7.7±6.7	19.5±10.5	3.8±4.0	52.3±10.6
<i>Subject Classifications at 69ms</i>										
Beach	47.2±6.4	10.8±5.5	18.7±6.8	4.3±2.7	7.7±5.0	1.9±1.9	1.9±1.9	2.5±2.4	2.6±2.3	2.3±2.3
Forest	5.2±3.6	63.5±6.4	11.3±6.0	2.2±2.5	5.1±4.5	6.8±4.4	3.2±2.9	2.7±2.6	0.0±0.0	0.0±0.0
Mtn	6.5±5.3	10.0±5.5	64.0±8.9	5.9±5.2	2.3±2.5	8.7±5.2	2.6±2.7	0.0±0.0	0.0±0.0	0.0±0.0
City	6.5±3.5	1.6±2.1	8.9±4.7	50.2±6.9	7.4±4.4	14.4±7.4	3.3±2.6	2.0±2.0	1.7±1.9	4.0±3.1
Farm	10.0±5.0	2.3±2.5	7.3±4.5	9.9±4.6	58.4±6.3	2.1±2.2	2.5±3.0	2.5±2.8	2.4±2.4	2.6±2.8
Street	3.3±2.7	8.9±4.6	1.7±1.8	13.6±5.2	10.7±4.5	47.5±5.9	1.9±1.9	6.7±4.1	1.8±2.2	3.9±3.2
Bath	3.5±2.3	0.0±0.0	1.8±1.9	3.8±2.4	2.0±2.3	1.6±1.8	44.3±5.3	5.3±3.4	22.2±7.3	15.7±5.7
Bedrm	2.6±2.6	0.0±0.0	4.9±3.6	2.8±2.7	0.0±0.0	3.0±3.0	6.6±3.8	60.3±5.8	2.2±2.4	17.5±6.8
Kitchn	4.6±3.6	5.1±3.7	1.8±1.9	4.3±3.3	0.0±0.0	9.4±4.6	12.6±5.0	6.6±4.5	50.9±6.0	4.6±3.8
Livingrm	2.0±2.2	2.4±2.3	1.8±2.2	0.0±0.0	5.9±4.5	0.0±0.0	5.1±3.6	20.2±8.1	5.0±3.4	57.7±7.3
<i>Texture Model Classifications</i>										
Beach	36.8±5.6	3.6±2.3	13.7±4.3	14.9±3.8	12.6±3.3	8.4±4.0	2.7±1.6	2.3±1.7	3.6±2.0	1.4±1.7
Forest	3.7±2.3	42.9±7.1	4.1±2.7	10.1±6.1	4.3±1.6	10.7±4.3	5.5±3.2	5.9±2.9	2.0±1.6	10.9±3.2
Mtn	16.3±3.9	6.1±3.1	24.6±6.9	14.8±5.5	15.2±4.6	8.8±4.2	3.3±2.5	4.0±2.3	2.4±1.7	4.7±3.0
City	7.6±2.7	9.2±3.4	6.9±3.7	34.3±5.6	5.8±2.1	12.8±4.1	5.4±2.3	6.2±3.2	4.6±3.1	7.1±2.3
Farm	13.8±4.8	4.5±2.3	8.3±5.1	10.9±3.9	28.8±5.3	11.0±4.4	4.5±2.7	6.1±3.3	7.6±2.5	4.5±2.3
Street	3.8±2.2	10.6±3.1	4.7±2.5	11.4±4.9	6.0±2.2	18.6±5.0	9.4±3.8	9.7±3.2	8.1±3.0	17.8±4.7
Bath	4.1±1.7	5.9±3.9	2.1±1.5	5.3±3.0	4.6±3.3	11.5±4.3	31.5±4.9	11.9±4.7	11.3±3.4	11.8±5.3
Bedrm	1.0±1.1	5.7±1.6	0.9±1.1	4.8±2.3	3.6±1.9	11.1±4.3	5.8±2.8	24.4±7.5	8.0±3.0	34.6±9.0
Kitchn	2.4±2.2	1.7±2.1	1.2±1.1	3.0±2.0	6.4±2.5	8.1±3.3	15.3±6.2	11.3±4.8	39.9±5.8	10.7±5.0
Livingrm	0.1±0.4	6.9±3.0	0.4±0.9	5.9±4.5	1.2±1.6	9.1±3.5	5.5±3.2	17.0±5.6	5.5±3.0	48.2±9.2

Table B.2. Superordinate-level confusion matrices for subjects and the texture model. Collapsing across the basic-level categories yields these matrices.

Scene Shown	Category Selected (%)		
	Natural / Outdoor	Man-made / Outdoor	Man-made / Indoor
<i>Subject Classifications at 37ms</i>			
Natural / Outdoor	48.32 \pm 4.60	20.64 \pm 3.11	31.05 \pm 4.39
Man-made / Outdoor	31.33 \pm 3.98	41.84 \pm 3.62	26.82 \pm 3.72
Man-made / Indoor	11.58 \pm 2.22	19.12 \pm 3.09	69.30 \pm 3.15
<i>Subject Classifications at 50ms</i>			
Natural / Outdoor	52.88 \pm 3.81	30.27 \pm 4.29	16.85 \pm 2.93
Man-made / Outdoor	25.44 \pm 3.66	56.46 \pm 3.79	18.10 \pm 3.57
Man-made / Indoor	8.84 \pm 1.87	17.54 \pm 2.67	73.62 \pm 3.03
<i>Subject Classifications at 62ms</i>			
Natural / Outdoor	75.71 \pm 5.45	10.28 \pm 4.05	14.00 \pm 4.57
Man-made / Outdoor	19.58 \pm 4.95	66.31 \pm 5.95	14.11 \pm 4.76
Man-made / Indoor	13.33 \pm 3.98	16.20 \pm 3.95	70.48 \pm 4.87
<i>Subject Classifications at 69ms</i>			
Natural / Outdoor	79.07 \pm 4.14	15.00 \pm 3.98	5.94 \pm 2.09
Man-made / Outdoor	16.85 \pm 3.37	71.37 \pm 4.00	11.77 \pm 2.39
Man-made / Indoor	7.62 \pm 1.92	8.19 \pm 2.37	84.18 \pm 2.84
<i>Texture Model Classifications</i>			
Natural / Outdoor	50.56 \pm 4.25	33.26 \pm 4.20	16.19 \pm 2.52
Man-made / Outdoor	23.14 \pm 3.15	46.54 \pm 3.38	30.33 \pm 3.26
Man-made / Indoor	8.12 \pm 1.51	18.69 \pm 3.32	73.18 \pm 3.10