02458 Cognitive Modelling

Bayesian Multisensory integration – Homework 3.3

Introduction

We are generally better at a perceptual task if we can use more than one sensory modality to perform it. One example is spatial localisation. We are much better at localising a sound if it is accompanied by a visual stimulus. This is due to a process of cross-modal perceptual integration, which is illustrated by the ventriloquist effect: it appears that a voice comes out of the dummy. The voice obviously comes out of the ventriloquist but due to the coordination of the dummy's mouth and the voice we experience an illusory displacement of the origin of the sound.

The strong fusion model has been very influential in modelling this and other audio-visual illusions. The model is based on Bayes' rule and some simplifying assumptions.

Let's denote the true azimuthal angle of the stimulus as S. As in signal detection theory and psychophysics we assume that the azimuthal angle is represented by an internal representation value, x_A. The internal representation is influenced by Gaussian noise centred on the true value, $x_A \sim \varphi(x_A|\mu_A, \sigma_A)$. When there is no visual stimulus and the observer performs the task based on hearing only. In terms of Bayes' rule the observer's response, R, reflects an estimate of the azimuthal angle, S, based on the internal representation value, x_A, which in the case of an uninformative prior can be written as:

$$P(R) = P(S|x_A) = \frac{P(x_A|S)P(S)}{P(x_A)} = \varphi(x_A|\mu_A, \sigma_A)$$

Similarly, when there is no auditory stimulus and the observer performs the task based only on a visual stimulus, the observer's estimate of S is based on an internal representation value, x_V.

When the observer is presented with both an auditory stimulus and a visual stimulus strong fusion asserts that the observer makes only a single estimate of the stimulus based on the auditory and visual internal representation values

$$P(R) = P(S|x_A, x_V) = \frac{P(x_A, x_V|S)P(S)}{P(x_A, x_V)}$$

A reasonable assumption often used in strong fusion models is that the Gaussian noise in the auditory system is independent from the Gaussian noise in the visual system. If we again assume that the observer as no prior knowledge of the location of the stimulus, the strong fusion model becomes

$$P(R) = P(S|x_A, x_V) = \frac{P(x_A|S)P(x_V|S)}{P(x_A, x_V)} = \frac{\varphi(x_A|\mu_A, \sigma_A)\varphi(x_V|\mu_V, \sigma_V)}{P(x_A, x_V)}$$

The next assumption is that the internal representation values are really amodal, meaning that the observer cannot tell if an internal representation value was caused by a visual stimulus or an auditory stimulus; x_A and x_V are just two values of x, hence

$$\frac{\varphi(x_A|\mu_A, \sigma_A)\varphi(x_V|\mu_V, \sigma_V)}{P(x_A, x_V)} = \frac{\varphi(x|\mu_A, \sigma_A)\varphi(x|\mu_V, \sigma_V)}{P(x)}$$

It turns out that the normalised product of two Gaussians is a Gaussian so that

$$\frac{\varphi(x|\mu_A, \sigma_A)\varphi(x|\mu_V, \sigma_V)}{P(x)} = \varphi(x|\mu_{AV}, \sigma_{AV})$$

where

$$\mu_{AV} = w\mu_A + (1 - w)\mu_V$$

$$w = \frac{\sigma_V^2}{\sigma_A^2 + \sigma_V^2}$$

$$\sigma_{AV}^2 = \frac{\sigma_A^2 \sigma_V^2}{\sigma_A^2 + \sigma_V^2}$$

Problems

An observer has completed 50 trials in an auditory, a visual and an audio-visual spatial localisation task. The data is in files xA, xV and xAV. Fit the strong fusion model to the data. Remember to use all the data to estimate the parameters.

What are your estimates of the free parameters?

What is the negative log likelihood of your fit?

Plot the estimated auditory, visual and audio-visual probability densities and the normalised histograms. Compare the probability densities and the histograms to assess the goodness-of-fit.

Another observer has completed the same experiment. This mind of this observer tends to wander. In 4% of the trials the observer zones out, forgets the task and guess more or less at random. In the remaining trials the observer behaves strictly according to the strong fusion model. How does the zoning out affect the goodness-of-fit?