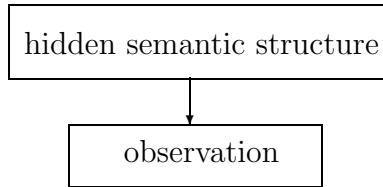


15 Interpreting senses (perception)

We use probabilistic models to build interpretation systems.

Generic model



Interpretation

$$\text{interp} = \arg \max_{\text{possible truth}} P(\text{possible truth} | \text{observation})$$

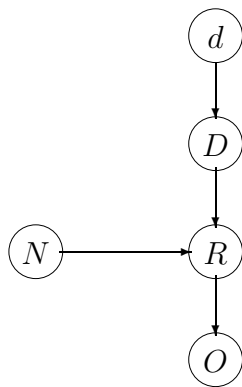
15.1 Toy example: Handwritten digit recognition

(This is Assignment 3!)

Given an image, recognize the digit

We consider two generative probabilistic models

Generative Model 1



Pick a digit $d \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 0\}$
(with uniform probability for each digit)

Get corresponding image template (a matrix)

$$D = \begin{bmatrix} d_{11} & \dots & d_{18} \\ \vdots & \ddots & \vdots \\ d_{81} & \dots & d_{88} \end{bmatrix} \in \{1, 2, 3, 4\}^{8 \times 8}$$

Generate a noise matrix

$$N = \begin{bmatrix} n_{11} & \dots & n_{18} \\ \vdots & \ddots & \vdots \\ n_{81} & \dots & n_{88} \end{bmatrix} \in \{-3, -2, -1, 0, 1, 2, 3\}^{8 \times 8}$$

Add noise matrix to template to generate raw image $r_{ij} = d_{ij} + n_{ij}$

$$R = \begin{bmatrix} r_{11} & \dots & r_{18} \\ \vdots & \ddots & \vdots \\ r_{81} & \dots & r_{88} \end{bmatrix} \in \{-2, -1, 0, 1, 2, 3, 4, 5, 6, 7\}^{8 \times 8}$$

Threshold raw image values to obtain final observed image

$$o_{ij} = \begin{cases} r_{ij} & \text{if } 1 \leq r_{ij} \leq 4 \\ 1 & \text{if } r_{ij} < 1 \\ 4 & \text{if } r_{ij} > 4 \end{cases}$$

$$O = \begin{bmatrix} o_{11} & \dots & o_{18} \\ \vdots & \ddots & \vdots \\ o_{81} & \dots & o_{88} \end{bmatrix} \in \{1, 2, 3, 4\}^{8 \times 8}$$

Note Once the digit d has been selected, each pixel is generated independently. That is, the pixels are conditionally independent *given* the digit identity d .

Recognition with Model 1

Given an image O , compute

$$\begin{aligned} d^* &= \arg \max_d P(d|O) \\ &= \arg \max_d \frac{P(dO)}{P(O)} \\ &= \arg \max_d P(dO) \\ &\quad \text{since } P(O) \text{ is a constant with respect to } d \end{aligned}$$

Now note that $P(dO)$ is a marginal

$$\begin{aligned}
P(dO) &= \sum_D \sum_N \sum_R P(dDNR O) \\
&= \sum_D \sum_N \sum_R P(d) P(DNR O|d) \\
&= \sum_D \sum_N \sum_R P(d) \prod_{i,j} P(d_{ij}, n_{ij}, r_{ij}, o_{ij}|d) \\
&\quad \text{since pixels are conditionally independent given } d \\
&= \sum_D \sum_N \sum_R P(d) \prod_{i,j} P(d_{ij}|d) P(n_{ij}) P(r_{ij}|d_{ij}, n_{ij}) P(o_{ij}|r_{ij}) \\
&\quad \text{by the Bayesian network structure behind each pixel}
\end{aligned}$$

But notice

$$\begin{aligned}
P(d_{ij}|d) &= \begin{cases} 1 & \text{if } d_{ij} = d_{ij}^{(d)} \\ 0 & \text{otherwise} \end{cases} \\
P(r_{ij}|d_{ij}, n_{ij}) &= \begin{cases} 1 & \text{if } r_{ij} = d_{ij} + n_{ij} \\ 0 & \text{otherwise} \end{cases} \\
P(o_{ij}|r_{ij}) &= \begin{cases} 1 & \text{if } r_{ij} = o_{ij} \\ & \text{or } r_{ij} < 1 \text{ and } o_{ij} = 1 \\ & \text{or } r_{ij} > 4 \text{ and } o_{ij} = 4 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

So combining the last two of these yields

$$\begin{aligned}
P(o_{ij}|d_{ij}, n_{ij}) &= \sum_{r_{ij}} P(o_{ij}, r_{ij}|d_{ij}, n_{ij}) \\
&= \sum_{r_{ij}} P(o_{ij}|r_{ij}, d_{ij}, n_{ij}) P(r_{ij}|d_{ij}, n_{ij}) \\
&= \sum_{r_{ij}} P(o_{ij}|r_{ij}) P(r_{ij}|d_{ij}, n_{ij}) \\
&\quad \text{by the Bayesian network structure} \\
&= \begin{cases} 1 & \text{if } d_{ij} + n_{ij} = o_{ij} \\ & \text{or } d_{ij} + n_{ij} < 1 \text{ and } o_{ij} = 1 \\ & \text{or } d_{ij} + n_{ij} > 4 \text{ and } o_{ij} = 4 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

Thus we have

$$\begin{aligned}
 P(dO) &= \sum_N P(d) \prod_{i,j} P(n_{ij}) P(o_{ij}|d_{ij}^{(d)}, n_{ij}) \\
 &= P(d) \sum_{n_{11}} \cdots \sum_{n_{88}} \prod_{i,j} P(n_{ij}) P(o_{ij}|d_{ij}^{(d)}, n_{ij}) \\
 &= P(d) \prod_{i,j} \sum_{n_{ij}} P(n_{ij}) P(o_{ij}|d_{ij}^{(d)}, n_{ij})
 \end{aligned}$$

distributing the sums over the products

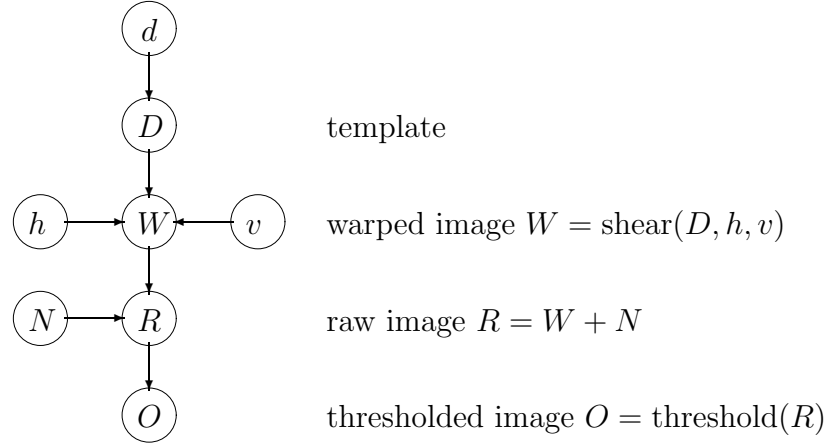
Now, to calculate conditional probabilities, note that for each $d = 0, 1, 2, \dots, 9$, we can just use

$$P(d|O) = \frac{P(dO)}{P(O)} = \frac{P(dO)}{\sum_{d'=0}^9 P(d'O)}$$

Generative Model 2

Add a global shear transformation over the whole image.

Generation



Here the shear variables h and v are independently chosen from $h \in \{0, 1, 2\}$ and $v \in \{-1, 0, 1\}$ respectively, and the shear operation is defined by

$$w_{ij} = \begin{cases} d_{t(j,i,v) t(i,j,h)} & \text{if both } 1 \leq t(j, i, v) \leq 8 \text{ and } 1 \leq t(i, j, h) \leq 8 \\ 1 & \text{otherwise} \end{cases}$$

where

$$\begin{aligned} t(j, i, v) &= i + \text{round}(v(5 - j)/4) \\ t(i, j, h) &= j + \text{round}(h(5 - i)/4) \end{aligned}$$

Recognition with Model 2

First note that pixels are no longer conditionally independent given d . They are conditionally independent given d , h and v . The pixels w_{ij} , r_{ij} and o_{ij} are deterministic functions

$$\begin{aligned} w_{ij} &= d_{t(j,i,v) t(i,j,h)} \quad \text{where } t(j, i, v) \text{ and } t(i, j, h) \text{ are as above} \\ r_{ij} &= n_{ij} + w_{ij} \\ o_{ij} &= \text{threshold}(n_{ij} + w_{ij}) \end{aligned}$$

Therefore, as with Model 1, we can simplify the Bayesian network by “collapsing” the deterministic functions into O and marginalizing out W and R , yielding

$$\begin{aligned}
P(dO) &= \sum_h \sum_v \sum_D \sum_N P(dhvDNO) \\
&= \sum_h \sum_v \sum_D \sum_N P(dhv) P(DNO|dhv) \\
&= \sum_h \sum_v \sum_D \sum_N P(dhv) \prod_{i,j} P(d_{ij}, n_{ij}, o_{ij} | d, h, v) \\
&\quad \text{since pixels are conditionally independent given } d, h \text{ and } v \\
&= \sum_h \sum_v \sum_D \sum_N P(d) P(h) P(v) \prod_{i,j} P(d_{ij}|d) P(n_{ij}) P(o_{ij}|d_{ij}, h, v, n_{ij}) \\
&\quad \text{by the Bayesian network structure} \\
&= \sum_h \sum_v \sum_N P(d) P(h) P(v) \prod_{i,j} P(n_{ij}) P(o_{ij}|d_{ij}^{(d)}, h, v, n_{ij}) \\
&\quad \text{where } P(o_{ij}|d_{ij}^{(d)}, h, v, n_{ij}) = \begin{cases} 1 & \text{if } o_{ij} = \text{threshold} \left(n_{ij} + d_{t(j,i,v) t(i,j,h)}^{(d)} \right) \\ 0 & \text{otherwise} \end{cases} \\
&= P(d) \sum_h P(h) \sum_v P(v) \sum_N \prod_{i,j} P(n_{ij}) P(o_{ij}|d_{ij}^{(d)}, h, v, n_{ij}) \\
&= P(d) \sum_h P(h) \sum_v P(v) \prod_{i,j} \sum_{n_{ij}} P(n_{ij}) P(o_{ij}|d_{ij}^{(d)}, h, v, n_{ij}) \\
&\quad \text{by distributing sums over products}
\end{aligned}$$

Conditional probabilities $P(d|O)$ can be calculated in the same way as for Model 1.

Readings

Russell and Norvig: Chapter 24

Dean, Allen, Aloimonos: Chapter 9