



Fig. 4.12: **Automatic localization of anatomy in 3D CT images.** (a) A coronal view of the abdomen of a patient in a CT scan. The bounding box of the right kidney is shown in orange. (b) Each voxel \mathbf{p} in the volume votes for the position of the six walls of the box via the relative displacements $d^R(\mathbf{p})$, $d^L(\mathbf{p})$, and so on.

For a voxel \mathbf{p} its feature vector $\mathbf{v}(\mathbf{p}) = (x_1, \dots, x_i, \dots, x_d) \in \mathbb{R}^d$ is a collection of differences:

$$x_i = \frac{1}{|\mathbf{B}_i|} \sum_{\mathbf{q} \in \mathbf{B}_i} J(\mathbf{q}). \quad (4.5)$$

where $J(\mathbf{p})$ denotes the density of the tissue in an element of volume at position \mathbf{p} as measured by the CT scanner (in calibrated Hounsfield Units). The 3D feature box \mathbf{B} (not to be confused with the output organ bounding box) is displaced from the reference point \mathbf{p} (see fig. 4.13a). Since for each reference pixel \mathbf{p} we can look at an infinite number of possible feature boxes ($\forall \mathbf{B} \in \mathbb{R}^6$) we have $d = \infty$.

During training we are given a database of CT scans which have been manually labelled with 3D boxes around organs of interest. A regression forest is trained to learn the association of voxel features and bounding box location. Training is achieved by maximizing a continuous information gain as in (4.1). Assuming multivariate Gaussian distributions at the nodes yields the already known form of continuous