Sparse representation has been successfully applied to visual tracking by finding the best candidate with a minimal reconstruction error using target templates. However, most sparse representation-based tracking methods only consider holistic rather than local appearance to discriminate between target and background regions, and hence may not perform well when target objects are heavily occluded. In this paper, we develop a simple yet robust tracking algorithm based on a coarse and fine structural local sparse appearance model. The proposed method exploits both partial and structural information of a target object based on sparse coding using the dictionary composed of patches from multiple target templates. The likelihood obtained by averaging and pooling operations exploits consistent appearance of object parts, thereby helping not only locate targets accurately but also handle partial occlusion. To update templates more accurately without introducing occluding regions, we introduce an occlusion detection scheme to account for pixels belonging to the target objects. The proposed method is evaluated on a large benchmark data set with three evaluation metrics. Experimental results demonstrate that the proposed tracking algorithm performs favorably against several state-of-the-art methods.