The recognition of events in video data is a subject of much current interest. In this paper, we address several issues related to this topic. The first one is overfitting when very large feature spaces are used and relatively small amounts of training data are available. The second is the use of a framework that can recognise events at different time scales, as standard Hidden Markov Model (HMM) do not model well long–term term temporal dependencies in the data. In this paper we propose a method combining Layered HMMs and an unsupervised low level clustering of the features to address these issues. Experiments conducted on the recognition task of different events in 8 rugby games demonstrates the potential of our approach with respect to standard HMM techniques coupled with a feature size reduction technique. While the current focus of this work focus is on events in sports videos, we believe the techniques shown here are general enough to be applied to other sources of data.