We propose the use of attentional cascades based on the DCT and motion information contained in an MPEG coded stream. An attentional cascade is a sequence of very efficient classifiers that reject a large number of negative candidate regions, while keeping all the positive candidates. Working directly on the compressed domain has two main advantages: computationally expensive features are already computed, and the stream is only partially decoded without the additional cost of full decompression, which will be reached by a very small number of the initial candidate regions. We have applied these concepts to skin color detection, as a pre-attentive filtering prior to face detection, and to text region detection with particular focus on license plates for vehicle identification. In both cases, a reduction of the number of candidate regions close to 95% is achieved, which turns into an enormous performance increase in video indexing processes.