

ABSTRACT

The present disclosure relates to the field of video learning. Provided are a video representation self-supervised contrastive learning method and apparatus. The method comprises : according to optical flow information corresponding to each video frame of a video clip, obtaining, by means of calculation, a motion amplitude diagram which corresponds to each video frame of the video clip; according to the motion amplitude diagram corresponding to each video frame of the video clip, determining motion information which corresponds to the video clip; and performing video representation self-supervised contrastive learning according to a video clip sequence, and the motion information corresponding to each video clip. Therefore, a motion-focused contrastive learning solution for video representation self-supervised learning is implemented, such that widely existing and very important motion information in a video is fully utilized during a learning process, thereby improving the performance of video representation self-supervised contrastive learning.