

Monocular video foreground/background segmentation by tracking spatial-color gaussian mixture models

Foreground/background segmentation is of great interest in many applications. It can be used in Video Conferencing for replacing the background of the speaker with a better environment, hence preventing the privacy of the speaker. It is used for tracking objects, motion detection, and various other image processing tasks.

The paper presents a new approach to segment foreground/background objects in a video. Instead of a stereo camera pair, this paper's method can be applied to videos captured just by monocular videos, though videos are assumed to be captured by static or hand-held monocular cameras filming large moving non-rigid foreground objects. The method is flexible to some extent since cameras are allowed to shake to somewhat extent and background objects can move too.

To solve the segmentation problem, first foreground and background are modelled separately using Spatio-color Gaussian Mixture models(SCGMM), which are built into Random Markov field energy function. This energy function is minimized by graph cut algorithm which leads to binary segmentation of the video frames.

The major problem in this process was that if the objects are moving rapidly across the frames or there are occlusions, then SCGMM models learns from previous frames are not usable for segmenting the current frames. To solve this problem, this paper introduces a 'foreground/background SCGMM joint tracking algorithm', which can propagate SCGMM models over the frames.

The drawback of the approach is that in the presence of occlusions or rapid motions, segmentation of objects is poor for subsequent some frames, though it gets better with time due to the joint tracking algorithm