**Problem statement**

We are going to recognize hand gestures from a video sequence. To recognize these gestures from a live video sequence, we first need to take out the hand region alone removing all the unwanted portions in the video sequence. After segmenting the hand region, we then count the fingers shown in the video sequence to instruct a robot based on the finger count. Thus, the entire problem could be solved using 2 simple steps -

1. Find and segment the hand region from the video sequence.
2. Count the number of fingers from the segmented hand region in the video sequence.

**Step 1.**

So, our first step to find the hand region from a video sequence involves three simple steps.

1. Background Subtraction
2. Motion Detection and Thresholding
3. Contour Extraction

**Step 2.**

**Four Intermediate Steps**

1. Find the convex hull of the segmented hand region (which is a contour) and compute the most extreme points in the convex hull (Extreme Top, Extreme Bottom, Extreme Left, Extreme Right).
2. Find the center of palm using these extremes points in the convex hull.
3. Using the palm’s center, construct a circle with the maximum Euclidean distance (between the palm’s center and the extreme points) as radius.
4. Perform bitwise AND operation between the thresholded hand image (frame) and the circular ROI (mask). This reveals the finger slices, which could further be used to calcualate the number of fingers shown.