Offline on Template Matching

Assignment: In this assignment, you are required to track a reference object in a video file (movie file) with ".MOV" extension.

There are three files here:

1. reference.jpg
2. input.MOV
3. output.MOV

The "input.MOV" movie file contains multiple frames. Each frame contains the reference image exactly once. Also, the reference image is present in the frames e*xactly,* meaning you do not have to employ any deformation based template matching here.

Your task is to read the video file, separate the frames, track the reference file in the frame, mark it with some marker and merge the frames back to a video file name "output.MOV". The given output file marks the reference with a red rectangle.

You have to do this entire job for the following methods:

1. Exhaustive search technique
2. 2D Logarithmic search
3. Hierarchical search

Exhaustive search is pretty straightforward, but for the other two, the location of the previous frame where the object was found must be used in searching the next frame. That is exactly what makes the search faster.

You also have to show the performance of these methods for different window sizes. A window [-p, +p] x [-p, +p] is defined by a single integer **p**. So you can vary the size of p and record some performance metric (say number of times entire frame is checked). Then for all the frames, you have a certain reading for a certain p. Taking average, you will get an estimation of that performance metric for a single value of p. Then, for another p, repeat the entire process. Finally, plot the estimations against p. Do this for all three methods.

For example, say there are 3 frames in the video. And with p = 2, Method X has to search the entire frame

a. 2 times for frame 1

b. 3 times for frame 2
c. 1 time for frame 3

Then for p = 2, X has the value: (1+2+3)/3 = 2. Similarly you have to calculate them for Methods Y and Z.
Naturally, you do not need to do this for the Exhaustive search because it does not depend on window size and searches the entire frame blindly. But for the other two, you must perform this.

So to sum up, your program should read “reference.jpg” and “input.MOV” ; and

Task1: output a file “output.MOV” with a proper marker

Task2: output sufficient numerical data on another file showing comparison amongst the three methods. A sample comparison:

P	Exhaustive	2D Log	Hierarchical
p1	x	y1	z1
p2	x	y2	z2
.	.	.	.
.	.	.	.
.	.	.	.

Guidelines: It is suggested that you finish task 1 completely first, and then move on to task 2. Even before starting task 1, try to work with two images. Search a large image and try to find a reference in it. When you are able to do so, then move on to dealing with the video. When you can deal with individual images, then dealing with a video should come to you naturally.

Coding: Use any language you like. But for submission, put your student ID in the source code files. Those using java should zip the entire project and upload it.

For video processing, you can use any library you like. But you must do the template matching codes yourselves.

Submission: Deadline is Monday, 6 AM of 8th Week. (24 April, 6 AM).