To solve the descripted problems, I devided that problems into smaller pieces and solved those in sequence as following diagram.

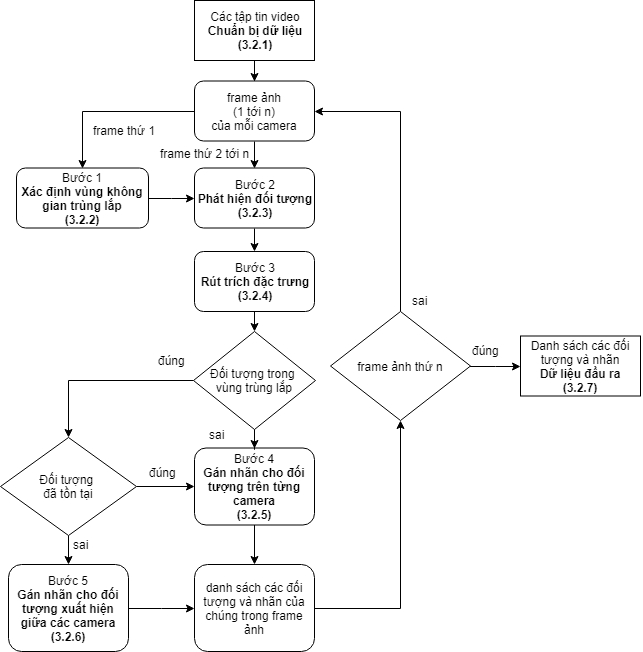


Diagram 1: The process of assigning label for moving objects in multiple camera.

1. prepared the data.

I prepared data from two sources, the dataset of [] and my experiments.

The input data for my proposed approach is two videos named “camera1.avi” and “camera2.avi”.

1. Determine the overlapped field of view.

We have two kind of overlapped field of view: two cameras were set in parallel and otherwise.

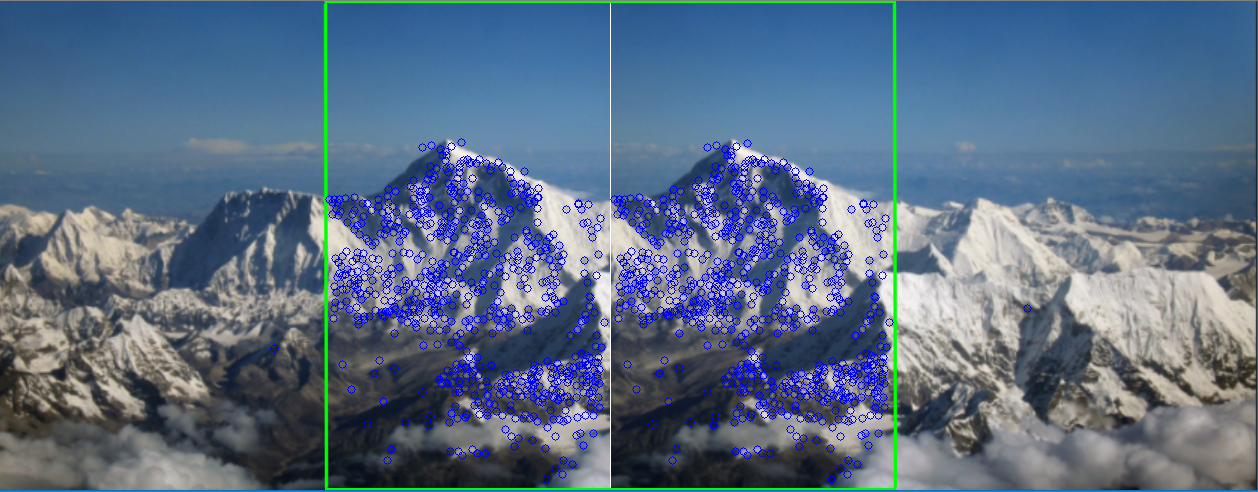
With two cameras were set in parallel:

+ Step 1: get the first frame of each input videos.

+ step 2: extract SIFT feature and find the landmark points of each frame.

+ Step 3: find all pair of landmark points between two frames.

+ Step 4: find the mapping function between those two frame and the set of those pairs is the overlapped fov. Figure 1 show the result of using sift feature to find the common view between two cameras.



**figure 3.2: using sift descriptor to find the common FOV[24].**

With two cameras weren’t set in parallel:

Look at the figure 2, we could see the common view of camera C1 and camera C2 is the quadrangle ABCD. The method to find the common FOV in this case is using a person move follow the edge of the C1’s view. When that person first emerged in C1, we mark it A and continue when that person disappeared in that way, I mark it B and with that, I could find C and D.

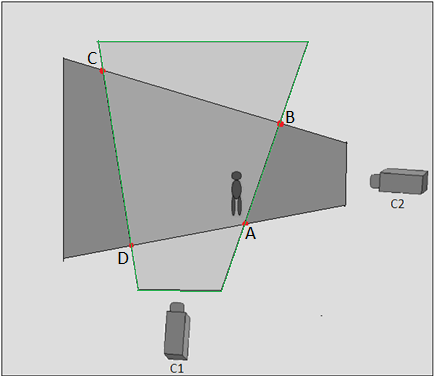


Figure 3.3: find the common FOV manually.

1. Detect objects

To handle the object detection problem, I use YOLO [1]. It’s a CNN trained in VOC 2007 and 2012 dataset. YOLO network has 24 Convolutional Layer and 2 final Full Convolutional Layer. The CL extract features from images such as HAAR, SIFT, HOG while the FCL predict the output of network. The output of YOLO network contains the position of detected objects (top, left, bottom, right) and the confidence that network recognized the objects are right.

1. Extract object’s features.

To help increase the accuracy of label object assigning process, I extract color diagram and shape features (hu’s invariant moment).

1. Assign label for objects in each camera.

To assign label for moving objects in each camera, I use Kalman Filter. The process of Kalman Filter is showed in figure 3.4



Figure 3.4: Kalman Filter circle.

In figure 3.4, Kalman Filter has two phases: predict phase and correct phase. In predict phase, Kalman Filter predict the value of objects in the next frame base on the set of its parameter. In correct phase, Kalman Filter got the measurement of those objects and update its parameter.

To increase the accuracy of Kalman Filter in predict objects, I use IoU to find the most correspond measurements of prediction values.

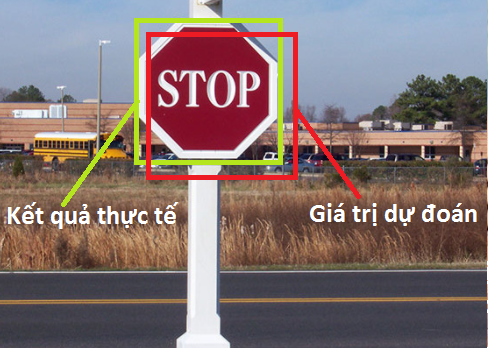


Figure 3.5: the predict value and its measurement.

In figure 3.5, I have two value of that traffic sign: the prediction value and the measurement value. IoU’s value is the ratio between the intersection and the union of the prediction value and the measurement value. The IoU computation is showed as figure 3.6:

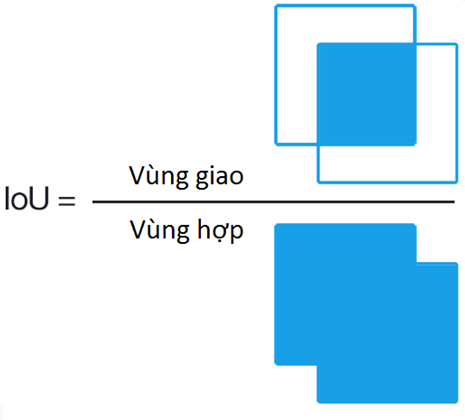


Figure 3.6: IoU computation.

The meaning of IoU represent for the similarity between the prediction value and the measurement value. The more IoU value, the much similarity it has when compare with the measurement. The figure 3.7 describe that:

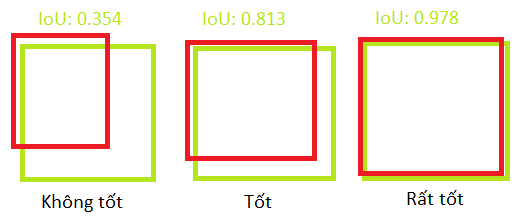


Figure 3.7: meaning of IoU.

1. Assign label for objects appear in common view
2. Export output.