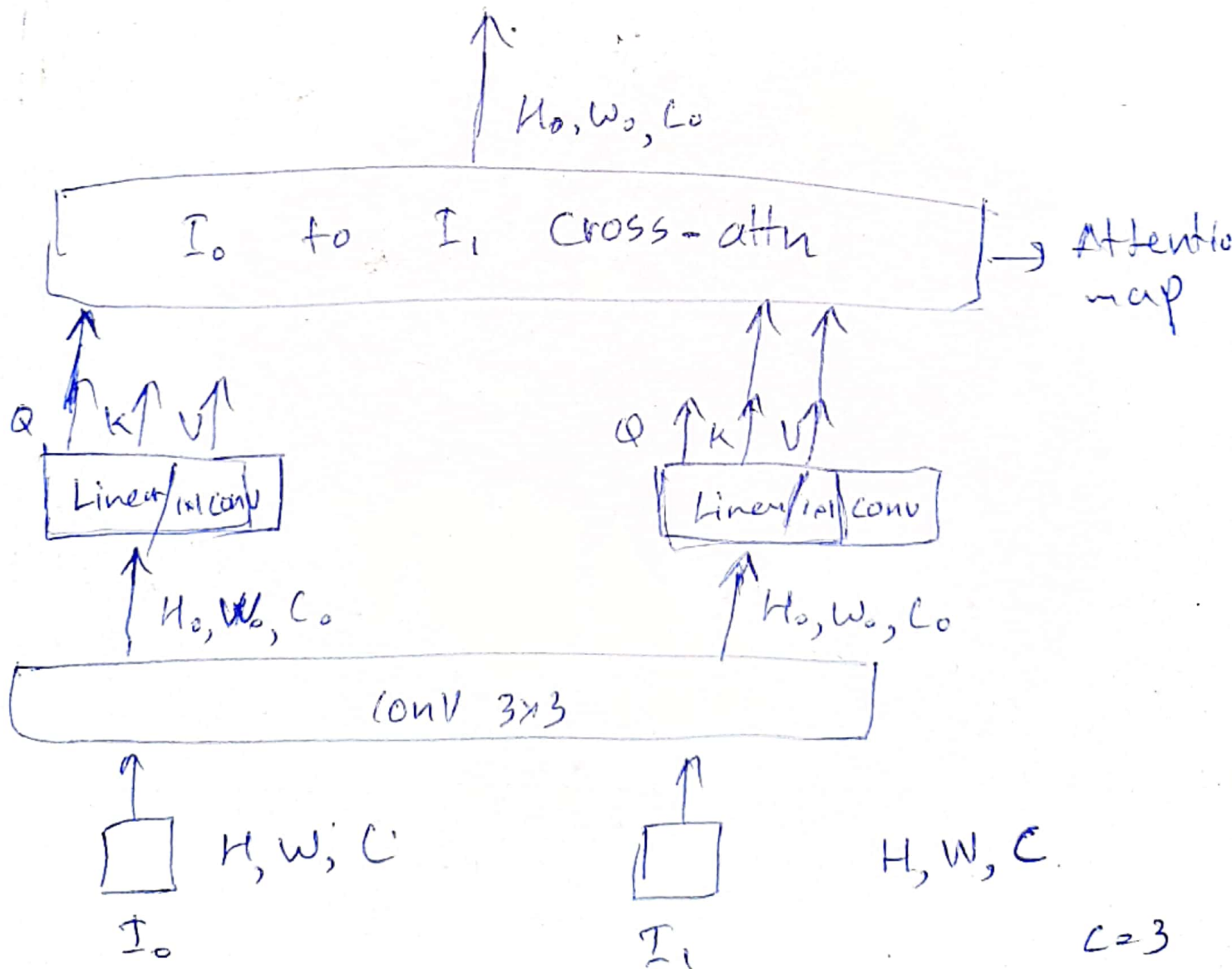
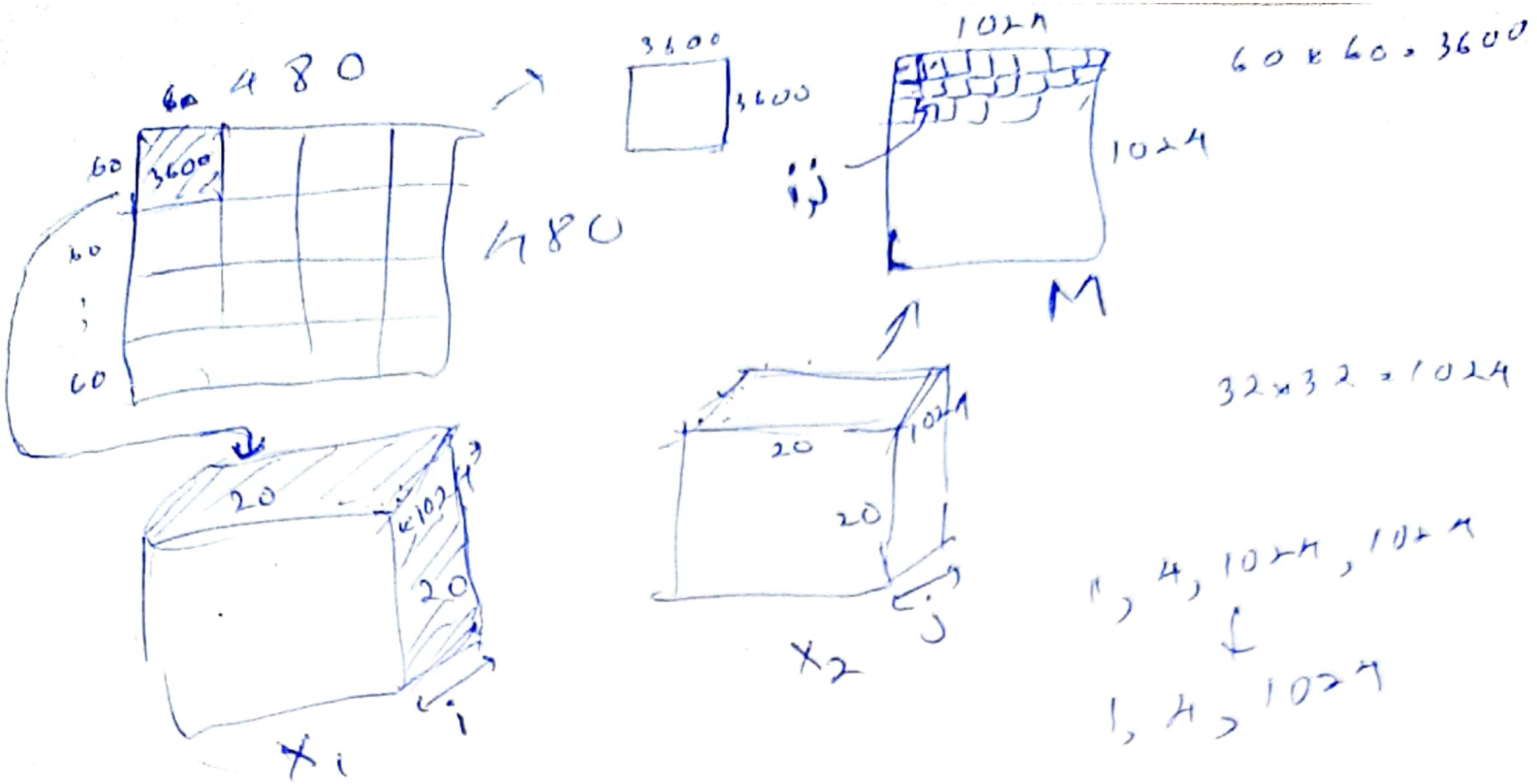


Attention of the conv features between two consecutive images.





- (ii) in M

1) select the indexes ~~where~~ where the attention is higher than a threshold.
- 2) Identify those feature maps from X_1 and X_2 which has more attention (more important features/ similar features) between two images. (more common context)
- 3) use the fused features and normal feature maps in Grad cam and compare the improvements after feature fusion across consequent frames.