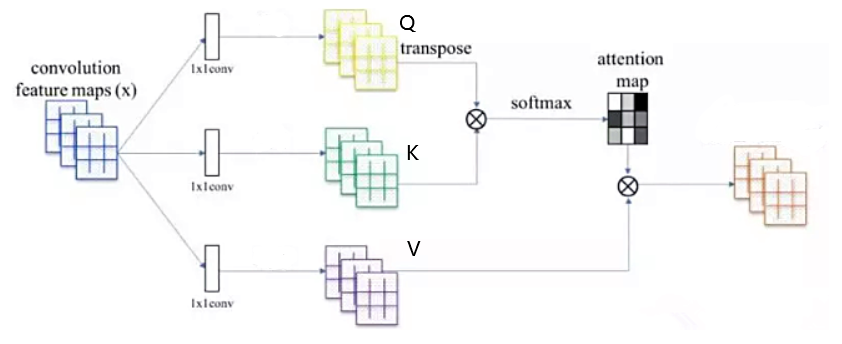
# 注意力机制

1.空间注意力：

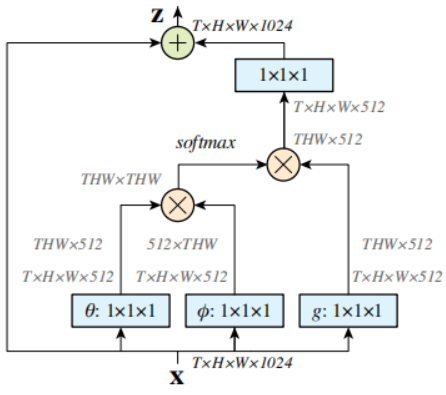
**1.1自注意力：Self-Attention**

自注意力计算时通常分为三步：

1. 第一步是将query和每个key进行相似度计算得到权重，常用的相似度函数有点积，拼接，感知机等；
2. 第二步一般是使用一个softmax函数对这些权重进行归一化，转换为注意力；
3. 第三步将权重和相应的键值value进行加权求和得到最后的attention。



**1.2 非局部注意力：**[**Non-local Attention**](https://link.zhihu.com/?target=http%3A//arxiv.org/abs/1711.07971)



(1) 首先对输入的feature map X 进行线性映射（1\*1\*1 卷积，来压缩通道数），然后得到θ，Φ，g特征；

(2)然后对θ，Φ进行相似度计算，对自相关特征以列或以行（具体看矩阵g 的形式而定） 进行Softmax 操作，得到0~1的权重，这里就是我们需要的Self-attention 系数；

(4) 最后将attention系数，对应乘回特征矩阵g 中，然后加上原输入的特征图，获得non-local block的输出。

**2.通道注意力**

通道域注意力类似于给每个通道上的特征图都施加一个权重，来代表该通道与关键信息的相关度的话，这个权重越大，则表示相关度越高。在神经网络中，越高的维度特征图尺寸越小，通道数越多，通道就代表了整个图像的特征信息。

2.1 SENet

