1、SIFT综述

尺度不变特征转换(Scale-invariant feature transform或SIFT)是一种电脑视觉的算法用来侦测与描述影像中的局部性特征，它在空间尺度中寻找极值点，并提取出其位置、尺度、旋转不变量，此算法由 David Lowe在1999年所发表，2004年完善总结。

其应用范围包含物体辨识、机器人地图感知与导航、影像缝合、3D模型建立、手势辨识、影像追踪和动作比对。

此算法有其专利，专利拥有者为英属哥伦比亚大学。

局部影像特征的描述与侦测可以帮助辨识物体，SIFT 特征是基于物体上的一些局部外观的兴趣点而与影像的大小和旋转无关。对于光线、噪声、些微视角改变的容忍度也相当高。基于这些特性，它们是高度显著而且相对容易撷取，在母数庞大的特征数据库中，很容易辨识物体而且鲜有误认。使用 SIFT特征描述对于部分物体遮蔽的侦测率也相当高，甚至只需要3个以上的SIFT物体特征就足以计算出位置与方位。在现今的电脑硬件速度下和小型的特征数据库条件下，辨识速度可接近即时运算。SIFT特征的信息量大，适合在海量数据库中快速准确匹配。

SIFT算法的特点有：

1. SIFT特征是图像的局部特征，其对旋转、尺度缩放、亮度变化保持不变性，对视角变化、仿射变换、噪声也保持一定程度的稳定性；

2. 独特性（Distinctiveness）好，信息量丰富，适用于在海量特征数据库中进行快速、准确的匹配；

3. 多量性，即使少数的几个物体也可以产生大量的SIFT特征向量；

4. 高速性，经优化的SIFT匹配算法甚至可以达到实时的要求；

5. 可扩展性，可以很方便的与其他形式的特征向量进行联合。

SIFT算法可以解决的问题：

目标的自身状态、场景所处的环境和成像器材的成像特性等因素影响图像配准/目标识别跟踪的性能。而SIFT算法在一定程度上可解决：

1. 目标的旋转、缩放、平移（RST）

2. 图像仿射/投影变换（视点viewpoint）

3. 光照影响（illumination）

4. 目标遮挡（occlusion）

5. 杂物场景（clutter）

6. 噪声

SIFT算法的实质是在不同的尺度空间上查找关键点(特征点)，并计算出关键点的方向。SIFT所查找到的关键点是一些十分突出，不会因光照，仿射变换和噪音等因素而变化的点，如角点、边缘点、暗区的亮点及亮区的暗点等。

Lowe将SIFT算法分解为如下四步：

1. 尺度空间极值检测：搜索所有尺度上的图像位置。通过高斯微分函数来识别潜在的对于尺度和旋转不变的兴趣点。

2. 关键点定位：在每个候选的位置上，通过一个拟合精细的模型来确定位置和尺度。关键点的选择依据于它们的稳定程度。

3. 方向确定：基于图像局部的梯度方向，分配给每个关键点位置一个或多个方向。所有后面的对图像数据的操作都相对于关键点的方向、尺度和位置进行变换，从而提供对于这些变换的不变性。

4. 关键点描述：在每个关键点周围的邻域内，在选定的尺度上测量图像局部的梯度。这些梯度被变换成一种表示，这种表示允许比较大的局部形状的变形和光照变化。

本文沿着Lowe的步骤，参考Rob Hess及Andrea Vedaldi源码，详解SIFT算法的实现过程。

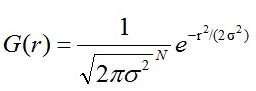
2、高斯模糊

SIFT算法是在不同的尺度空间上查找关键点，而尺度空间的获取需要使用高斯模糊来实现，Lindeberg等人已证明高斯卷积核是实现尺度变换的唯一变换核，并且是唯一的线性核。本节先介绍高斯模糊算法。

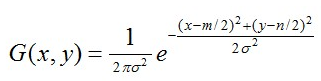
2.1二维高斯函数

高斯模糊是一种图像滤波器，它使用正态分布(高斯函数)计算模糊模板，并使用该模板与原图像做卷积运算，达到模糊图像的目的。

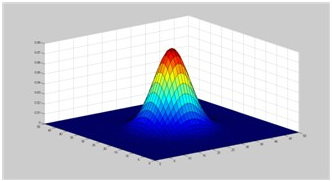
N维空间正态分布方程为：

 （1-1）

其中，是正态分布的标准差，值越大，图像越模糊(平滑)。r为模糊半径，模糊半径是指模板元素到模板中心的距离。如二维模板大小为m\*n，则模板上的元素(x,y)对应的高斯计算公式为：

 （1-2）

   在二维空间中，这个公式生成的曲面的等高线是从中心开始呈正态分布的同心圆，如图2.1所示。分布不为零的像素组成的卷积矩阵与原始图像做变换。每个像素的值都是周围相邻像素值的加权平均。原始像素的值有最大的高斯分布值，所以有最大的权重，相邻像素随着距离原始像素越来越远，其权重也越来越小。这样进行模糊处理比其它的均衡模糊滤波器更高地保留了边缘效果。



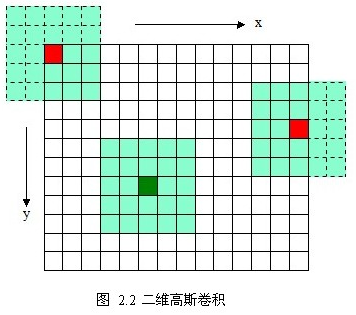
理论上来讲，图像中每点的分布都不为零，这也就是说每个像素的计算都需要包含整幅图像。在实际应用中，在计算高斯函数的离散近似时，在大概3σ距离之外的像素都可以看作不起作用，这些像素的计算也就可以忽略。通常，图像处理程序只需要计算的矩阵就可以保证相关像素影响。

2.2 图像的二维高斯模糊

根据σ的值，计算出高斯模板矩阵的大小()，使用公式(1-2)计算高斯模板矩阵的值，与原图像做卷积，即可获得原图像的平滑(高斯模糊)图像。为了确保模板矩阵中的元素在[0,1]之间，需将模板矩阵归一化。5\*5的高斯模板如表2.1所示。



下图是5\*5的高斯模板卷积计算示意图。高斯模板是中心对称的。



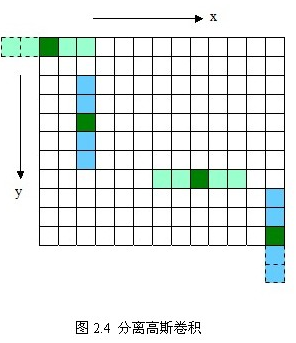


2.3分离高斯模糊

如图2.3所示，使用二维的高斯模板达到了模糊图像的目的，但是会因模板矩阵的关系而造成边缘图像缺失(2.3 b,c)，σ越大，缺失像素越多,丢弃模板会造成黑边(2.3 d)。更重要的是当σ变大时，高斯模板(高斯核)和卷积运算量将大幅度提高。根据高斯函数的可分离性，可对二维高斯模糊函数进行改进。

高斯函数的可分离性是指使用二维矩阵变换得到的效果也可以通过在水平方向进行一维高斯矩阵变换加上竖直方向的一维高斯矩阵变换得到。从计算的角度来看，这是一项有用的特性，因为这样只需要次计算，而二维不可分的矩阵则需要次计算，其中，m,n为高斯矩阵的维数，M,N为二维图像的维数。

另外，两次一维的高斯卷积将消除二维高斯矩阵所产生的边缘。(关于消除边缘的论述如下图2.4所示， 对用模板矩阵超出边界的部分——虚线框，将不做卷积计算。如图2.4中x方向的第一个模板1\*5，将退化成1\*3的模板，只在图像之内的部分做卷积。)



3、尺度空间极值检测

尺度空间使用高斯金字塔表示。Tony Lindeberg指出尺度规范化的LoG(Laplacion of Gaussian)算子具有真正的尺度不变性，Lowe使用高斯差分金字塔近似LoG算子，在尺度空间检测稳定的关键点。

3.1 尺度空间理论

尺度空间(scale space)思想最早是由Iijima于1962年提出的，后经witkin和Koenderink等人的推广逐渐得到关注，在计算机视觉邻域使用广泛。

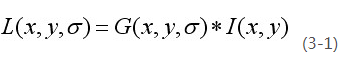
尺度空间理论的基本思想是：在图像信息处理模型中引入一个被视为尺度的参数，通过连续变化尺度参数获得多尺度下的尺度空间表示序列，对这些序列进行尺度空间主轮廓的提取，并以该主轮廓作为一种特征向量，实现边缘、角点检测和不同分辨率上的特征提取等。

尺度空间方法将传统的单尺度图像信息处理技术纳入尺度不断变化的动态分析框架中，更容易获取图像的本质特征。尺度空间中各尺度图像的模糊程度逐渐变大，能够模拟人在距离目标由近到远时目标在视网膜上的形成过程。

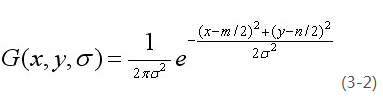
尺度空间满足视觉不变性。该不变性的视觉解释如下：当我们用眼睛观察物体时，一方面当物体所处背景的光照条件变化时，视网膜感知图像的亮度水平和对比度是不同的，因此要求尺度空间算子对图像的分析不受图像的灰度水平和对比度变化的影响，即满足灰度不变性和对比度不变性。另一方面，相对于某一固定坐标系，当观察者和物体之间的相对位置变化时，视网膜所感知的图像的位置、大小、角度和形状是不同的，因此要求尺度空间算子对图像的分析和图像的位置、大小、角度以及仿射变换无关，即满足平移不变性、尺度不变性、欧几里德不变性以及仿射不变性。

3.2 尺度空间的表示

一个图像的尺度空间，定义为一个变化尺度的高斯函数与原图像的卷积。



其中，\*表示卷积运算，



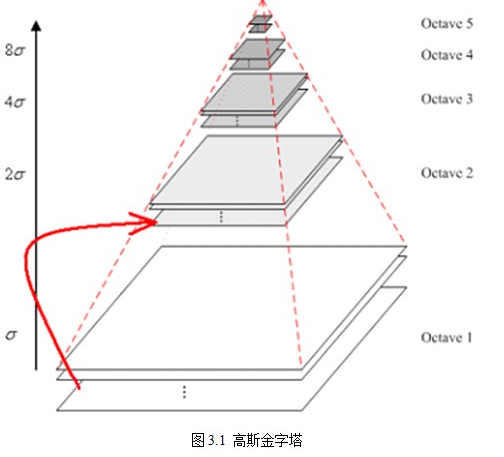
与公式(1-2)相同，m，n表示高斯模板的维度(由确定)。(x, y)代表图像的像素位置。σ是尺度空间因子，值越小表示图像被平滑的越少，相应的尺度也就越小。大尺度对应于图像的概貌特征，小尺度对应于图像的细节特征。

3.3 高斯金字塔的构建

尺度空间在实现时使用高斯金字塔表示，高斯金字塔的构建分为两部分：

1. 对图像做不同尺度的高斯模糊；

2. 对图像做降采样(隔点采样)。



图像的金字塔模型是指，将原始图像不断降阶采样，得到一系列大小不一的图像，由大到小，从下到上构成的塔状模型。原图像为金子塔的第一层，每次降采样所得到的新图像为金字塔的一层(每层一张图像)，每个金字塔共n层。金字塔的层数根据图像的原始大小和塔顶图像的大小共同决定，其计算公式如下：



其中M，N为原图像的大小,t为塔顶图像的最小维数的对数值。如，对于大小为512\*512的图像，金字塔上各层图像的大小如表3.1所示，当塔顶图像为4\*4时，n=7，当塔顶图像为2\*2时，n=8。



为了让尺度体现其连续性，高斯金字塔在简单降采样的基础上加上了高斯滤波。如图3.1所示，将图像金字塔每层的一张图像使用不同参数做高斯模糊，使得金字塔的每层含有多张高斯模糊图像，将金字塔每层多张图像合称为一组(Octave)，金字塔每层只有一组图像，组数和金字塔层数相等，使用公式(3-3)计算，每组含有多张(也叫层Interval)图像。另外，降采样时，高斯金字塔上一组图像的初始图像(底层图像)是由前一组图像的倒数第三张图像隔点采样得到的。

注：由于组内的多张图像按层次叠放，因此组内的多张图像也称做多层，为避免与金字塔层的概念混淆，本文以下内容中，若不特别说明是金字塔层数，层一般指组内各层图像。

注：如3.4节所示，为了在每组中检测S个尺度的极值点，则DOG金字塔每组需S+2层图像，而DOG金字塔由高斯金字塔相邻两层相减得到，则高斯金字塔每组需S+3层图像，实际计算时S在3到5之间。取S=3时，假定高斯金字塔存储索引如下：

第0组(即第-1组)：  0 1  2  3  4   5

第1组：            6 7  8  9  10 11

第2组：            ？

则第2组第一张图片根据第一组中索引为9的图片降采样得到，其它类似。

3.4 高斯差分金字塔

2002年Mikolajczyk在详细的实验比较中发现尺度归一化的高斯拉普拉斯函数的极大值和极小值同其它的特征提取函数，例如：梯度，Hessian或Harris角特征比较，能够产生最稳定的图像特征。

而Lindeberg早在1994年就发现高斯差分函数（Difference of Gaussian ，简称DOG算子）与尺度归一化的高斯拉普拉斯函数非常近似。其中和的关系可以从如下公式推导得到：



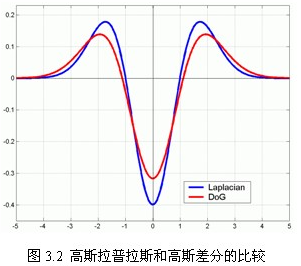
利用差分近似代替微分，则有：



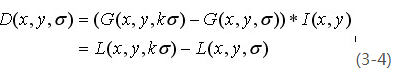
  因此有



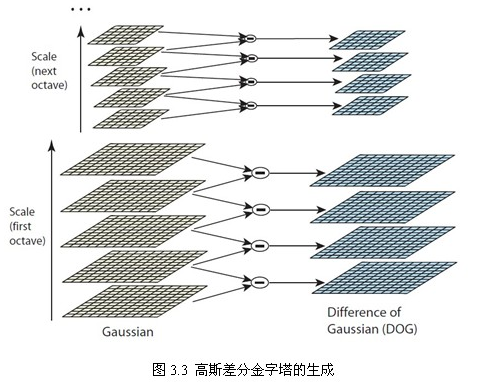
其中k-1是个常数，并不影响极值点位置的求取。



如图3.2所示，红色曲线表示的是高斯差分算子，而蓝色曲线表示的是高斯拉普拉斯算子。Lowe使用更高效的高斯差分算子代替拉普拉斯算子进行极值检测，如下：

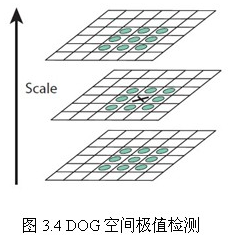


在实际计算时，使用高斯金字塔每组中相邻上下两层图像相减，得到高斯差分图像，如图3.3所示，进行极值检测。



3.5 空间极值点检测(关键点的初步探查)

关键点是由DOG空间的局部极值点组成的，关键点的初步探查是通过同一组内各DoG相邻两层图像之间比较完成的。为了寻找DoG函数的极值点，每一个像素点要和它所有的相邻点比较，看其是否比它的图像域和尺度域的相邻点大或者小。如图3.4所示，中间的检测点和它同尺度的8个相邻点和上下相邻尺度对应的9×2个点共26个点比较，以确保在尺度空间和二维图像空间都检测到极值点。



由于要在相邻尺度进行比较，如图3.3右侧每组含4层的高斯差分金子塔，只能在中间两层中进行两个尺度的极值点检测，其它尺度则只能在不同组中进行。为了在每组中检测S个尺度的极值点，则DOG金字塔每组需S+2层图像，而DOG金字塔由高斯金字塔相邻两层相减得到，则高斯金字塔每组需S+3层图像，实际计算时S在3到5之间。

当然这样产生的极值点并不全都是稳定的特征点，因为某些极值点响应较弱，而且DOG算子会产生较强的边缘响应。

3.6 构建尺度空间需确定的参数

σ—尺度空间坐标

    O—组(octave)数

    S— 组内层数

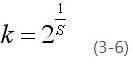
在上述尺度空间中，O和S，σ的关系如下：

 (3-5)

其中σ0是基准层尺度，o为组octave的索引，s为组内层的索引。关键点的尺度坐标就是按关键点所在的组和组内的层，利用公式(3-5)计算而来。

在最开始建立高斯金字塔时，要预先模糊输入图像来作为第0个组的第0层的图像，这时相当于丢弃了最高的空域的采样率。因此通常的做法是先将图像的尺度扩大一倍来生成第-1组。我们假定初始的输入图像为了抗击混淆现象，已经对其进行的高斯模糊，如果输入图像的尺寸用双线性插值扩大一倍，那么相当于。

取式(3-4)中的k为组内总层数的倒数，即



   在构建高斯金字塔时，组内每层的尺度坐标按如下公式计算：

(3-7)

其中初始尺度σ0，lowe取，s为组内的层索引，不同组相同层的组内尺度坐标相同。组内下一层图像是由前一层图像按进行高斯模糊所得。式(3-7)用于一次生成组内不同尺度的高斯图像，而在计算组内某一层图像的尺度时，直接使用如下公式进行计算：

 (3-8)

该组内尺度在方向分配和特征描述时确定采样窗口的大小。

由上，式(3-4)可记为

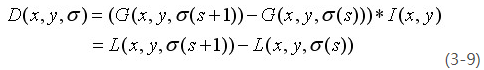
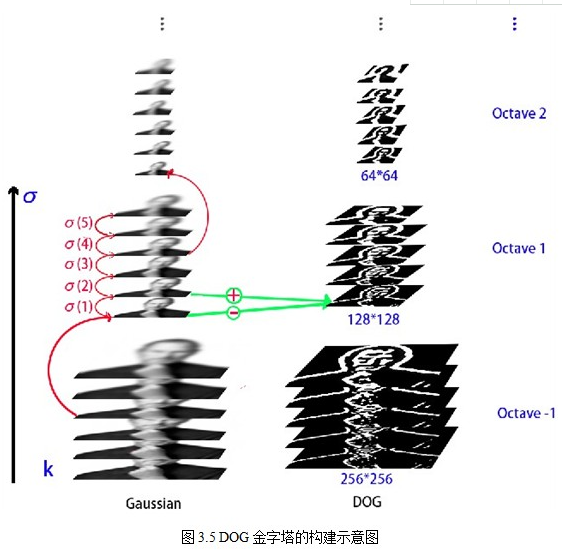


图3.5为构建DOG金字塔的示意图，原图采用128\*128的jobs图像，扩大一倍后构建金字塔。

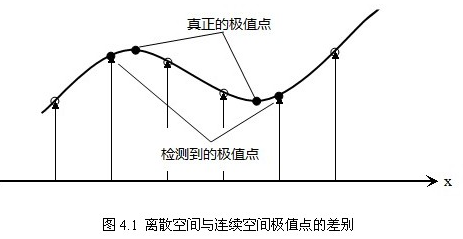


4、关键点定位

以上方法检测到的极值点是离散空间的极值点，以下通过拟合三维二次函数来精确确定关键点的位置和尺度，同时去除低对比度的关键点和不稳定的边缘响应点(因为DoG算子会产生较强的边缘响应)，以增强匹配稳定性、提高抗噪声能力。

4.1关键点的精确定位

离散空间的极值点并不是真正的极值点，图4.1显示了二维函数离散空间得到的极值点与连续空间极值点的差别。利用已知的离散空间点插值得到的连续空间极值点的方法叫做子像素插值（Sub-pixel Interpolation）。



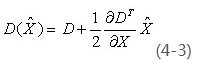
为了提高关键点的稳定性，需要对尺度空间DoG函数进行曲线拟合。利用DoG函数在尺度空间的Taylor展开式(拟合函数)为：

 (4-1)

其中，。求导并让方程等于零，可以得到极值点的偏移量为：

 (4-2)

对应极值点，方程的值为：

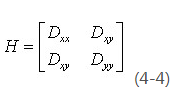


其中,代表相对插值中心的偏移量，当它在任一维度上的偏移量大于0.5时（即x或y或σ），意味着插值中心已经偏移到它的邻近点上，所以必须改变当前关键点的位置。同时在新的位置上反复插值直到收敛；也有可能超出所设定的迭代次数或者超出图像边界的范围，此时这样的点应该删除，在Lowe中进行了5次迭代。另外，过小的点易受噪声的干扰而变得不稳定，所以将小于某个经验值(Lowe论文中使用0.03，Rob Hess等人实现时使用0.04/S)的极值点删除。同时，在此过程中获取特征点的精确位置(原位置加上拟合的偏移量)以及尺度()。

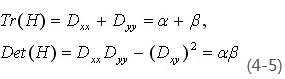
4.2消除边缘响应

一个定义不好的高斯差分算子的极值在横跨边缘的地方有较大的主曲率，而在垂直边缘的方向有较小的主曲率。

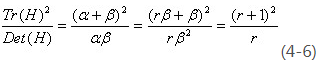
DOG算子会产生较强的边缘响应，需要剔除不稳定的边缘响应点。获取特征点处的Hessian矩阵，主曲率通过一个2x2 的Hessian矩阵H求出：



H的特征值α和β代表x和y方向的梯度，

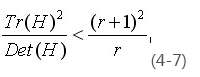


表示矩阵H对角线元素之和，表示矩阵H的行列式。假设是α较大的特征值，而是β较小的特征值，令，则



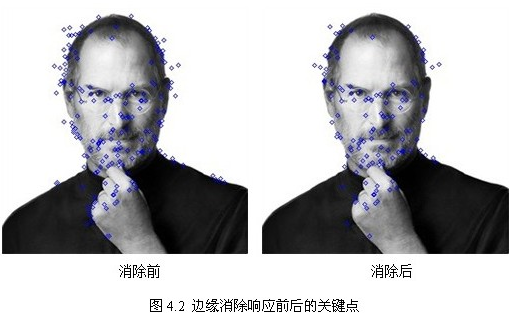
导数由采样点相邻差估计得到,在下一节中说明。

D的主曲率和H的特征值成正比，令为α最大特征值，β为最小的特征值，则公式的值在两个特征值相等时最小，随着的增大而增大。值越大，说明两个特征值的比值越大，即在某一个方向的梯度值越大，而在另一个方向的梯度值越小，而边缘恰恰就是这种情况。所以为了剔除边缘响应点，需要让该比值小于一定的阈值，因此，为了检测主曲率是否在某域值r下，只需检测



式(4-7)成立时将关键点保留，反之剔除。

在Lowe的文章中，取r＝10。图4.2右侧为消除边缘响应后的关键点分布图。



4.3有限差分法求导

有限差分法以变量离散取值后对应的函数值来近似微分方程中独立变量的连续取值。在有限差分方法中，我们放弃了微分方程中独立变量可以取连续值的特征，而关注独立变量离散取值后对应的函数值。但是从原则上说，这种方法仍然可以达到任意满意的计算精度。因为方程的连续数值解可以通过减小独立变量离散取值的间格，或者通过离散点上的函数值插值计算来近似得到。这种方法是随着计算机的诞生和应用而发展起来的。其计算格式和程序的设计都比较直观和简单，因而，它在计算数学中使用广泛。

有限差分法的具体操作分为两个部分：

1. 用差分代替微分方程中的微分，将连续变化的变量离散化，从而得到差分方程组的数学形式；

2. 求解差分方程组。

一个函数在x点上的一阶和二阶微商，可以近似地用它所临近的两点上的函数值的差分来表示。如对一个单变量函数f(x)，x为定义在区间[a,b]上的连续变量，以步长将区间[a,b]离散化，我们会得到一系列节点，



然后求出f(x)在这些点上的近似值。显然步长h越小，近似解的精度就越好。与节点相邻的节点有和，所以在节点处可构造如下形式的差值：

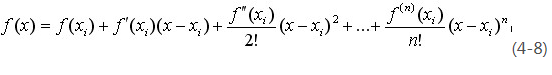
 节点的一阶向前差分

节点的一阶向后差分

节点的一阶中心差分

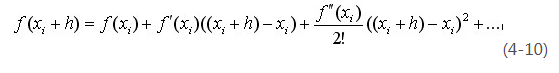
本文使用中心差分法利用泰勒展开式求解第四节所使用的导数，现做如下推导。

函数f(x)在处的泰勒展开式为：

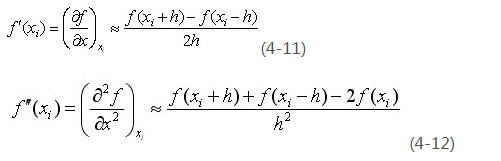


则，

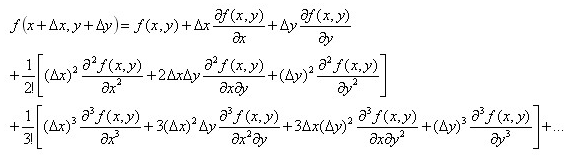




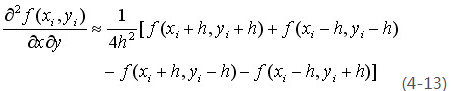
忽略h平方之后的项，联立式(4-9)，(4-10)解方程组得：



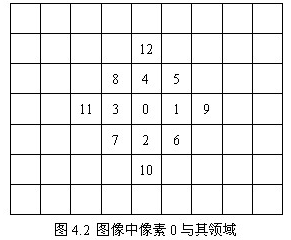
二元函数的泰勒展开式如下：



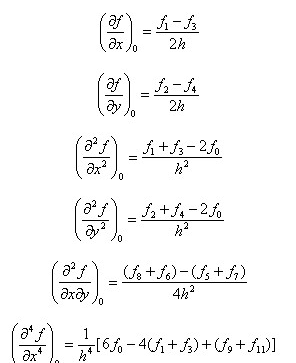
将展开后忽略次要项联立解方程得二维混合偏导如下：



综上，推导了4.1,4.2遇到的所有导数计算。同理，利用多元泰勒展开式，可得任意偏导的近似差分表示。

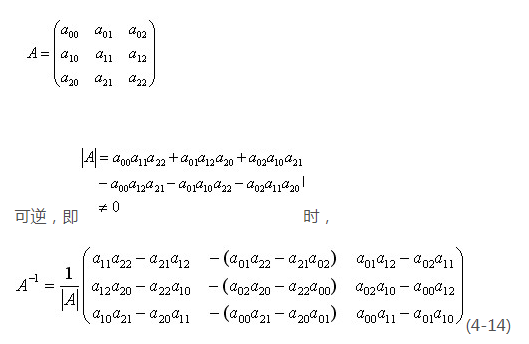


在图像处理中，取h=1，在图4.2所示的图像中，将像素0的基本中点导数公式整理如下：



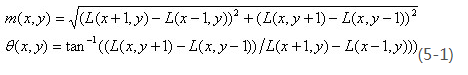
4.4 三阶矩阵求逆公式

高阶矩阵的求逆算法主要有归一法和消元法两种，现将三阶矩阵求逆公式总结如下：



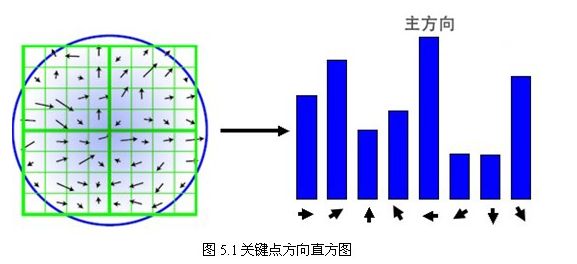
5、关键点方向分配

为了使描述符具有旋转不变性，需要利用图像的局部特征为给每一个关键点分配一个基准方向。使用图像梯度的方法求取局部结构的稳定方向。对于在DOG金字塔中检测出的关键点点，采集其所在高斯金字塔图像3σ邻域窗口内像素的梯度和方向分布特征。梯度的模值和方向如下：



L为关键点所在的尺度空间值，按Lowe的建议，梯度的模值m(x,y)按的高斯分布加成，按尺度采样的3σ原则，邻域窗口半径为。

在完成关键点的梯度计算后，使用直方图统计邻域内像素的梯度和方向。梯度直方图将0~360度的方向范围分为36个柱(bins)，其中每柱10度。如图5.1所示，直方图的峰值方向代表了关键点的主方向，(为简化，图中只画了八个方向的直方图)。



方向直方图的峰值则代表了该特征点处邻域梯度的方向，以直方图中最大值作为该关键点的主方向。为了增强匹配的鲁棒性，只保留峰值大于主方向峰值80％的方向作为该关键点的辅方向。因此，对于同一梯度值的多个峰值的关键点位置，在相同位置和尺度将会有多个关键点被创建但方向不同。仅有15％的关键点被赋予多个方向，但可以明显的提高关键点匹配的稳定性。实际编程实现中，就是把该关键点复制成多份关键点，并将方向值分别赋给这些复制后的关键点，并且，离散的梯度方向直方图要进行插值拟合处理，来求得更精确的方向角度值，检测结果如图5.2所示。

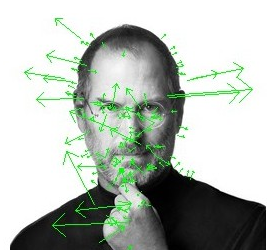


图5.2 shift 特征

至此，将检测出的含有位置、尺度和方向的关键点即是该图像的SIFT特征点。

6、关键点特征描述

通过以上步骤，对于每一个关键点，拥有三个信息：位置、尺度以及方向。接下来就是为每个关键点建立一个描述符，用一组向量将这个关键点描述出来，使其不随各种变化而改变，比如光照变化、视角变化等等。这个描述子不但包括关键点，也包含关键点周围对其有贡献的像素点，并且描述符应该有较高的独特性，以便于提高特征点正确匹配的概率。

SIFT描述子是关键点邻域高斯图像梯度统计结果的一种表示。通过对关键点周围图像区域分块，计算块内梯度直方图，生成具有独特性的向量，这个向量是该区域图像信息的一种抽象，具有唯一性。

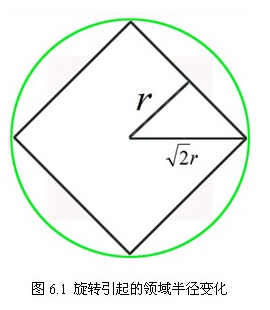
Lowe建议描述子使用在关键点尺度空间内4\*4的窗口中计算的8个方向的梯度信息，共4\*4\*8=128维向量表征。表示步骤如下：

1. 确定计算描述子所需的图像区域

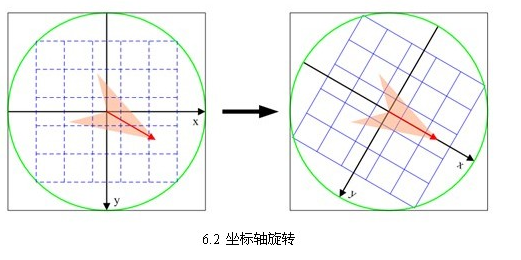
特征描述子与特征点所在的尺度有关，因此，对梯度的求取应在特征点对应的高斯图像上进行。将关键点附近的邻域划分为d\*d(Lowe建议d=4)个子区域，每个子区域做为一个种子点，每个种子点有8个方向。每个子区域的大小与关键点方向分配时相同，即每个区域有个子像素，为每个子区域分配边长为的矩形区域进行采样(实际用边长为的矩形区域即可包含，但由式(3-8)，不大，为了简化计算取其边长为，并且采样点宜多不宜少)。考虑到实际计算时，需要采用双线性插值，所需图像窗口边长为。在考虑到旋转因素(方便下一步将坐标轴旋转到关键点的方向)，如下图6.1所示，实际计算所需的图像区域半径为：



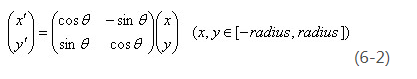
计算结果四舍五入取整。



2. 将坐标轴旋转为关键点的方向，以确保旋转不变性，如6.2所示。



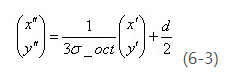
旋转后邻域内采样点的新坐标为：



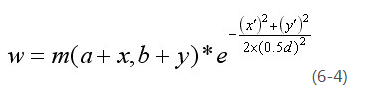
3. 将邻域内的采样点分配到对应的子区域内，将子区域内的梯度值分配到8个方向上，计算其权值。

旋转后的采样点坐标在半径为radius的圆内被分配到的子区域，计算影响子区域的采样点的梯度和方向，分配到8个方向上。

旋转后的采样点落在子区域的下标为

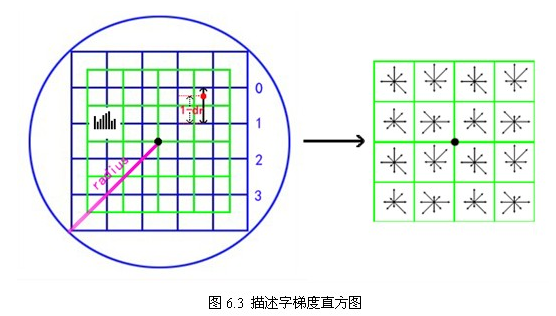


Lowe建议子区域的像素的梯度大小按的高斯加权计算，即



其中a，b为关键点在高斯金字塔图像中的位置坐标。

4. 插值计算每个种子点八个方向的梯度。

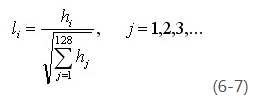


如图6.3所示，将由式(6-3)所得采样点在子区域中的下标(图中蓝色窗口内红色点)线性插值，计算其对每个种子点的贡献。如图中的红色点，落在第0行和第1行之间，对这两行都有贡献。对第0行第3列种子点的贡献因子为dr，对第1行第3列的贡献因子为1-dr，同理，对邻近两列的贡献因子为dc和1-dc，对邻近两个方向的贡献因子为do和1-do。则最终累加在每个方向上的梯度大小为：



其中k，m，n为0或为1。

5. 如上统计的4\*4\*8=128个梯度信息即为该关键点的特征向量。特征向量形成后，为了去除光照变化的影响，需要对它们进行归一化处理，对于图像灰度值整体漂移，图像各点的梯度是邻域像素相减得到，所以也能去除。得到的描述子向量为，归一化后的特征向量为则

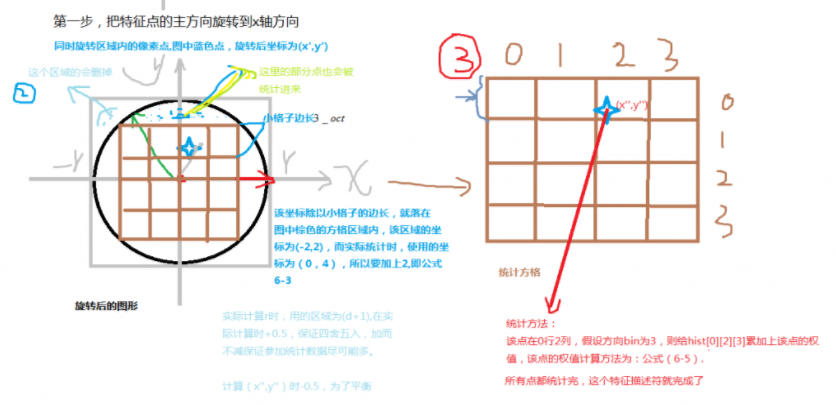
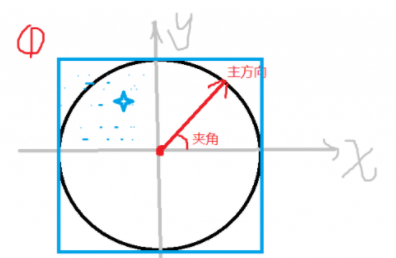


6. 描述子向量门限。非线性光照，相机饱和度变化对造成某些方向的梯度值过大，而对方向的影响微弱。因此设置门限值(向量归一化后，一般取0.2)截断较大的梯度值。然后，再进行一次归一化处理，提高特征的鉴别性。

7. 按特征点的尺度对特征描述向量进行排序。

至此，SIFT特征描述向量生成。

描述向量这块不好理解，我画了个草图，供参考：



7、SIFT的缺点

SIFT在图像的不变特征提取方面拥有无与伦比的优势，但并不完美，仍然存在：

1. 实时性不高。

2. 有时特征点较少。

3. 对边缘光滑的目标无法准确提取特征点。

等缺点，如下图7.1所示，对模糊的图像和边缘平滑的图像，检测出的特征点过少，对圆更是无能为力。近来不断有人改进，其中最著名的有SURF和CSIFT。

8、总结

本人研究SIFT算法一月有余，鉴于相关知识的缺失，尺度空间技术和差分近似求导曾困我良久。Lowe在论文中对细节提之甚少，甚至只字未提，给实现带来了很大困难。经过多方查阅，实现，总结成此文。自认为是到目前为止，关于SIFT算法最为详尽的资料，现分享给你，望批评指正。