

## 应用快速高斯变换估计图像互信息

蒋 鸿, 胡永祥, 朱文球

(湖南工业大学 计算机与通信学院, 湖南 株洲 412008)

(jzhong\_504@163.com)

**摘 要:** 针对基于核密度估计的图像互信息估计法运算量很大的问题, 提出了一种快速互信息估计算法。该算法选用高斯函数作为核密度估计法的核函数, 然后用快速高斯变换加快运算速度, 并用  $k$  中心聚类算法改进了原快速高斯变换中数据分类方法。新算法能在线性时间复杂度内估计出图像互信息, 采用医学图像的实验证实了算法的性能。

**关键词:** 互信息; 快速高斯变换; 核密度估计;  $k$  中心聚类; 图像配准

**中图分类号:** TP391.41 **文献标志码:** A

## Estimating mutual information of image by fast Gaussian transform

JIANG Hong, HU Yong-xiang, ZHU Wen-qiu

(Institute of Computer & Communication, Hunan University of Technology, Zhuzhou Hunan 412008, China)

**Abstract:** A new fast algorithm was presented to accelerate the computation of mutual information of images based on kernel density estimate. This algorithm adopted gauss function as kernel function, and then used fast Gaussian transform to reduce time complexity and improved the fast Gaussian transform by  $k$ -center clustering. The new algorithm can evaluate the mutual information in linear time complexity and experiments show it is very efficient.

**Key words:** mutual information; fast Gaussian transform; kernel density estimate;  $k$ -center clustering; image registration

### 0 引言

1995年, 文献[1, 2]首先提出用图像互信息作为图像配准的相似性测度, 随后众多研究者对它进行了深入的研究并提出了很多改进算法<sup>[3-8]</sup>。互信息是信息论的重要概念, 用来衡量两个随机变量的相关程度, 定义为:

$$I(X, Y) = \int \mu(x, y) \log \frac{\mu(x, y)}{\mu_x(x) \mu_y(y)} dx dy \quad (1)$$

其中  $\mu(x, y)$  为两个随机变量  $X, Y$  的联合概率密度,  $\mu_x(x) = \int \mu(x, y) dx$  和  $\mu_y(y) = \int \mu(x, y) dy$  为随机变量  $X$  和  $Y$  的边缘概率密度。在图像配准中, 将图像像素看作某个随机变量的采样, 当图像配准时, 两个图像的相关性最大, 互信息的值也就最大; 反之, 图像越不匹配, 图像互信息的值就越小。图像互信息测度可以用于单模态之间、多模态之间的图像配准, 也可以用于刚性和非刚性图像配准<sup>[9]</sup>, 是应用最多、最广的相似性测度。

由互信息的定义可知, 估计图像互信息需要估计单个图像的概率密度和两个图像的联合概率密度。常用的图像概率密度估计为核密度估计法<sup>[10-12]</sup>, 它不需要预先规定概率密度函数的形式且能得到平滑的估计函数, 一般形式为:

$$\hat{p}(x_j | S) = \sum_{i=1}^N w_i K(x_j - s(i)); j = 1, \dots, M \quad (2)$$

其中  $K(\cdot)$  称为核函数,  $w_i$  为加权系数 (通常取  $w_i = 1/N$ ),  $S = \{s(i), i = 1, \dots, N\}$  为给定的采样点集。容易看出, 核密度估计法的时间复杂度为  $O(MN)$ 。图像配准是一个迭代求解过程, 每次迭代都需要重新估计图像互信息。如果直接用核密度估计法来估计图像互信息, 图像配准过程耗时很长, 难以实际

应用。因此, 研究图像互信息的快速估计算法成为图像配准技术的重要课题。

快速估计图像互信息的关键在于加快核密度估计。目前已存在许多核密度估计的快速方法, 如利用  $k$  最近邻搜索算法<sup>[13]</sup>、分支界限法<sup>[14]</sup>、双树算法<sup>[15-16]</sup>等。核密度估计通常采用高斯函数作为核函数, 而快速高斯变换<sup>[17-20]</sup>能以  $O(M + N)$  的时间复杂度估算  $N$  个高斯函数的和在  $M$  个离散点上的值, 非常适合用来加快核密度估计。本文采用均值为零、方差为  $\sigma$  的高斯函数作为核密度估计法的核函数, 然后用快速高斯变换加快减少运算量, 并用  $k$  中心聚类法改进原快速高斯变换的空间数据划分方法, 从而得到估计图像互信息的快速估计算法。

### 1 快速高斯变换

快速高斯变换<sup>[17-20]</sup>用于快速估计如下形式的高斯函数加权求和:

$$G(t_i) = \sum_{j=1}^N f_j e^{-\frac{(t_i - s_j)^2}{\sigma^2}}; i = 1, \dots, M \quad (3)$$

其中  $f_j$  是加权系数,  $\{s_j\}_{j=1, \dots, N}$  是高斯函数的中心, 叫作源点,  $\sigma$  为高斯函数的带宽,  $\{t_i\}_{i=1, \dots, M}$  为目标点, 在这些点上估计  $N$  个高斯函数和的值。

定义 Hermite 函数  $h_n(x) = (-1)^n \frac{d^n}{dx^n} (e^{-x^2})$ , 高斯函数

的 Hermite 展开为:

$$e^{-(t-s)^2} = \sum_{n=0}^{\infty} \frac{s^n}{n!} h_n(t)$$

进行平移和尺度变换后有:

收稿日期: 2007-09-11; 修回日期: 2007-12-17。

作者简介: 蒋鸿 (1976-), 女, 湖南株洲人, 讲师, 硕士研究生, 主要研究方向: 智能信息处理、计算机视觉; 胡永祥 (1973-), 男, 湖南安化人, 讲师, 硕士, 主要研究方向: 医学图像处理、模式识别; 朱文球 (1969-), 男, 湖南攸县人, 副教授, 硕士, 主要研究方向: 图像处理、模式识别。

$$e^{-(t-t_0)^2/\sigma} = \sum_{n=0}^{\infty} \frac{1}{n!} \left( \frac{t-t_0}{\sqrt{\sigma}} \right)^n h_n \left( \frac{t-t_0}{\sqrt{\sigma}} \right) \quad (4)$$

根据式(3)、(4)得:

$$G(t_i) = \sum_{n=0}^{\infty} \left[ \sum_{j=1}^N f_j h_n \left( \frac{s_j - t_0}{\sqrt{\sigma}} \right) \right] \frac{1}{n!} \left( \frac{t_i - t_0}{\sqrt{\sigma}} \right)^n; i = 1, \dots, M \quad (5)$$

可以看出,截取前 $p$ 项后式(5)的时间复杂度为 $O(Np + Mp)$ 。由于级数快速衰减,通常 $p \ll \min(N, M)$ ,所以运算量大大地减少了。

将多变量的高斯函数看作多个单变量的高斯函数的乘积,可以将一维情况推广至高维,即:

$$G(t_i) = \sum_{n=0}^{\infty} \prod_{k=1}^d \left[ \sum_{j=1}^N f_j h_n \left( \frac{(s_j)_k - (t_0)_k}{\sqrt{\sigma}} \right) \right] \frac{1}{n!} \left( \frac{(t_i)_k - (t_0)_k}{\sqrt{\sigma}} \right)^n \quad (6)$$

同样可以看出,截取 $p$ 项后式(6)的时间复杂度为 $O((M + N)p^d)$ 。

## 2 空间数据划分

当源点 $s_i$ 和目标点 $t_i$ 相距较大时,高斯函数快速衰减,因此式(6)的计算并不需要涉及全部 $N$ 个源点。原快速高斯变换使用均匀的网格划分数据空间,只有与目标点相邻的 $n$ 个子空间的数据参与计算。然而,这种数据空间划分的方法在高维空间中会产生大量的子空间并且其中大部分子空间为空。假设在一个10维空间中,每一维分成10等分,则共有 $10^{10}$ 子空间,这使得数据的存储和处理将耗费大量的时间。

为克服上述问题,我们将数据划分问题看作 $k$ 中心聚类问题, $k$ 值一般由用户预先指定。该问题可描述为:在任意空间中给定 $N$ 个数据,且预定 $k$ 个类,将 $N$ 个点分配到中心为 $c_1, \dots, c_k$ 的 $S_1, \dots, S_k$ 中,使得 $\max_{v \in S_i} \|v - c_i\|$ 具有最小值。这是一个NP难问题。文献[21]提出了一个非常简单的近似算法,即:最近点聚类法。该算法为:

1) 任意选择一点 $v_0$ 作为第一个类的中心,并将其加入到类中心集 $C$ 中;

2) 对于 $i = 1, \dots, k$ ,在每次循环中计算每个点到类中心集 $C$ 的距离: $d_i(v, C) = \min_{c \in C} \|v - c\|$ ,令 $v_i$ 是与点集 $C$ 最远的点,将其加入集合 $C$ 。

3) 分配每个点到与其最近的类。

算法的时间复杂度为 $O(Nk)$ 。文献[22]将这个算法的时间复杂度改进为 $O(M \log k)$ ,并且证明了这个算法在代数决策树模型下是最优的。2001年,文献[23]利用hashing和随机采样技术将算法进一步改进为 $O(N)$ 。

## 3 估计图像互信息

估计图像互信息首先要将两个图像的像素构成二维向量集<sup>[1-2]</sup>或高维向量集<sup>[5-8]</sup>,然后用核密度估计法得到向量集的概率密度函数(联合概率密度)。在核密度估计法中,我们用均值为0的高斯函数作为核函数,且将高维高斯函数看作多个一维高斯函数相乘,即核函数为:

$$K(x) = \prod_{i=1}^d \frac{1}{\sqrt{2\sigma}} e^{-x_i^2/\sigma^2} \quad (7)$$

其中 $d$ 表示空间的维数, $x_i$ 表示 $d$ 维向量的第 $i$ 部分。这样核密度估计的问题转变成求多个高斯函数和的问题,可以采用快

速高斯变换减少运算量。接着,利用联合概率密度求出边缘概率密度。最后,求得图像互信息。综上所述,图像互信息的快速估计算法如下:

1) 设图像大小为 $N$ ,数据划分的类数为 $K$ ,概率密度的估计点个数为 $M$ 。给定图像 $A$ 和 $B$ ,指定邻域个数为 $m$ ,将两个图像对应位置处的像素 $A_{ij}$ 及 $m$ 个邻域 $B_{ij}$ 及 $m$ 个邻域一起构成一个 $2(m+1)$ 维向量集;

2) 用最远点聚类算法对向量集进行划分;

3) 用核密度估计法求出向量集的概率密度函数,并用快速高斯变换加快运算速度;

4) 根据3)的联合概率密度求出边缘概率密度;

5) 计算出图像互信息。

容易看出,算法的第1)、4)步的时间复杂度为 $O(N)$ ;第2)步为 $O(NK)$ ,当 $K$ 较大时可选用时间复杂度为 $O(M \log K)$ 或 $O(N)$ 的算法;第3)步为 $O(M + N)$ ;第5)步为 $O(1)$ 。因此,总的时间复杂度为 $O(M + N)$ 或 $O(NK)$ 。另外,在第1)步中 $m$ 等于0就是求文献[1-2]提出的互信息,不等于0则求出的是文献[5-8]所提出的高维互信息。

## 4 实验

我们采用大脑的断层扫描图像来进行实验,如图1所示。两幅图像取自大脑的同一位置,其中一个为CT图像,另一个为MRI图像,图像大小为 $256 \times 256$ ,并且这两个图像已配准。实验采用Matlab编程,在256 MB内存,1.3 GHz主频的电脑上运行。

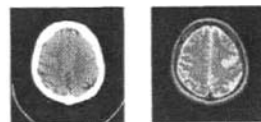


图1 实验图像

### 实验一 估计图像的二维互信息

比较利用新算法(kFGT, k-center clustering FGT)、核密度估计法(KDE)、原快速高斯变换(FGT)估计图像互信息的性能,实验结果如表1所示。实验中设置概率密度的估计点数为 $M = 128$ ,数据划分的类数 $k = 64$ ,而原快速高斯变换的子空间数为 $16 \times 16$ 。为了测试各种大小图像下的运算效果,我们将图1中的图像进行缩放,缩放时的插值算法采用双线性插值。

表1 二维互信息估计

图像大小	KDE	FGT	kFGT
64 × 64	1.22	0.031	0.030
128 × 128	4.97	0.103	0.091
256 × 256	19.56	0.430	0.370
512 × 512	78.46	1.405	1.342
1024 × 1024	314.96	5.874	5.187

从表1可以看出:1)基于FGT和kFGT的互信息估计方法所需时间远小于直接用KDE方法所需的时间,与图像大小基本成线性关系。这说明运用快速高斯变换估计图像互信息大大地减少了运算量,并且具有线性时间复杂度;2)基于FGT和kFGT的互信息估计方法所需时间相差很小。这是由于图像数据的取值区间为 $[0, 255]$ ,我们用基于FGT方法估计时将各种大小图像的数据子空间数固定,这使得用 $k$ 中心聚类

法进行数据空间划分与用均匀网格划分的方法相比没有体现出优势,这说明在低维空间两个方法的性能基本一致。

#### 实验二 估计图像的高维互信息

二维互信息没有充分利用图像固有的空间信息,文献[5-8]对其进行了扩充,利用像素邻域、梯度等来构成高维互信息。为进一步验证算法的性能,利用像素的1,2,3个邻域像素,构成4,6,8维互信息,分别用FGT和kFGT方法来估计,实验结果如表2所示。实验参数设置为 $M = 64 \times d, k = 32 \times d$ ,FGT方法中数据子空间数变为 $16^d, d = 4, 6, 8$ 。

表2 高维互信息估计

图像大小	FGT-4	KFGT-4	FGT-6	KFGT-6	FGT-8	KFGT-8
64×64	0.405	0.150	4.492	1.101	199.290	5.498
128×128	1.018	0.380	11.769	2.404	637.730	13.419
256×256	4.262	1.282	49.430	7.821	—	43.382
512×512	14.908	4.627	222.460	27.780	—	158.650
1024×1024	61.164	17.535	956.580	115.730	—	618.730

由表2可见,在4,6,8维数据空间中基于kFGT估计方法所需时间远远小于用基于FGT估计方法的时间,且随着维数的增加两个方法所需的时间也增大,但差距越来越大,不过在相同维空间仍保持与图像大小的线性关系。这主要是由于FGT方法随着维数的增加,用均匀网格法所划分的子空间成几何级数增加,造成子空间的存储和处理耗费大量时间,以致于在8维空间时对256×256大小以上的图像我们的电脑不能处理,而kFGT方法的分类数仅与维数成线性关系,所以耗费时间较少。同时应注意到,在8维数据空间中,即使对于256×256大小的图像估计互信息的时间也需要大约43s,这对于需迭代求解的图像配准来说是令人难以忍受的。

## 5 结语

为降低基于核密度估计法的图像互信息估计的时间复杂度,提出了一种应用快速高斯变换的快速估计算法,并用最远点聚类算法改进原快速高斯变换中空间数据划分法。采用MRI图像和CT图像进行了二维、高维互信息估计的实验。实验结果表明新算法能在线性时间复杂度内快速估计出图像互信息,且在高维互信息估计中,所需时间远小于原快速高斯变换所需时间。新算法的优越性使应用核密度估计法估计连续的图像互信息成为可能,有利于最优化算法精确地估计互信息梯度,快速、准确地找到相似性测度的最优值以取得高精度的配准结果。

#### 参考文献:

- [1] VIOLA P, WELLS W. Alignment by maximization of mutual information[C]// Proceedings of the 5th International Conference on Computer Vision, Boston: MA, 1995: 16-23.
- [2] COLLIGNON A, MAES F, VANDERMEULEN D, et al. Automated multimodality image registration using information theory[C]// Proceedings of the Information Processing in Medical Imaging Conference, Dordrecht: Kluwer Academic Publishers, 1995: 263-274.
- [3] STUDHOLME C, HILL D L G, HAWKES D J. An overlap invariant entropy measures of 3D medical image alignment[J]. Pattern Recognition, 1999, 32(1): 71-86.
- [4] PLUIM J P, MAINTZ J, VIERGEVER M A. Image registration by maximization of combined mutual information and gradient information[J]. IEEE Transactions on Medical Imaging, 2000, 19(8): 809-814.
- [5] RUECKERT D, CLARKSON M J, HILL D L G, et al. Non-rigid registration using higher-order mutual information[C]// Proceedings of SPIE Medical Imaging 2000. San Diego: IEEE, 2000, 438-447.
- [6] SABUNCU M R, RAMADGE P J. Spatial information in entropy-based image registration[C]// Proceedings of 2nd Workshop in Biomedical Image Registration (WBIR-03). Philadelphia: IEEE, 2003: 132-141.
- [7] HOLDEN M, GRIFFIN N I D, HILL D L G. Multi-dimensional mutual information image similarity metrics based on derivatives of linear scale space[C]// Proceedings of the APRS Workshop on Digital Image Computing. [S. l.]: IEEE, 2005: 55-60.
- [8] 卢振泰, 陈武凡. 基于共生互信息量的医学图像配准[J]. 计算机学报, 2007, 30(6): 1022-1027.
- [9] AGOSTINO E D, MAES F. A viscous fluid model for multimodal non-rigid image registration using mutual information[J]. Medical Image Analysis, 2003, 7(4): 565-575.
- [10] ROSENBLATT M. Remarks on some nonparametric estimates of a density function[J]. Annals of Mathematical Statistics, 1956, 27(6): 832-837.
- [11] PARZEN E. On estimation of a probability density function and model[J]. Annals of Mathematical Statistics, 1962, 33(8): 1065-1076.
- [12] VIOLA P A. Alignment by maximization of mutual information[D]. Massachusetts Institute of Technology—Artificial Intelligence Laboratory, 1995.
- [13] POSTAIRE J G, VASSEUR C. A fast algorithm for nonparametric probability density estimation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1982, 4(6): 663-666.
- [14] JEON B, LANDGREBE D A. Fast Parzen density estimation using clustering-based branch and bound[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(9): 950-954.
- [15] GRAY A, MOORE A. Nonparametric density estimation: Toward computational tractability[C]// SIAM International Conference on Data Mining, 2003: 301-316.
- [16] GRAY A G, MOORE A W. 'N-Body' problems in statistical learning[C]// Advances in Neural Information Processing Systems 4, 2000: 521-527.
- [17] GREENGARD L, STRAIN J. The fast Gauss transform[J]. SIAM J SCI Statistics and Computing, 1991, 12(1): 79-94.
- [18] STRAIN J. The fast Gauss transform with variable scales[J]. SIAM J SCI Statistics and Computing, 1991, 12(5): 1131-1139.
- [19] GREENGARD L, SUN X. A new version of the fast Gauss transform[C]// Proceedings of the International Congress of Mathematicians, III(Extra Vol. III). [S. l.]: IEEE, 1998: 575-584.
- [20] YANG CHANG-JIANG, DURAISWAMI R, GUMEROV N, et al. Improved fast Gauss transform and efficient kernel density estimation[C]// Proceedings of the Ninth IEEE International Conference on Computer Vision. Washington, DC: IEEE Computer Society, 2003: 464-471.
- [21] GONZALEZ T. Clustering to minimize and maximum intercluster distance[J]. Theoretical Computer Science, 1985, 38: 293-306.
- [22] FEDER T, GREENE D H. Optimal algorithms for approximate clustering[C]// Proceedings of the twentieth annual ACM symposium on Theory of computing. New York: ACM, 1988: 434-444.
- [23] HAR-PELED S. Clustering motion[C]// Proceedings of the 42nd IEEE symposium on Foundations of Computer Science. Washington, DC: IEEE Computer Society, 2001: 84-93.