

# 基于最大化对齐度的多模态图像自动配准<sup>1</sup>

王东峰 邹谋炎

(中国科学院电子学研究所 北京 100080)

**摘要** 该文讨论图像配准的广义定义并提出一种新的图像配准准则——对齐度。对齐度能够正确地反映两幅不同灰度属性的图像是否配准,并且将图像配准归结为对齐度的最大化。新准则可以用来配准多模态图像包括灰度属性差异很大的图像。该文应用多分辨率方法和 Powell 方向族搜索法来最大化对齐度。因此,用该方法可以自动地处理图像配准问题,不需要定义地标点或特征点。多模态医学图像和遥感图像的自动配准的例子说明了该文方法的有效性。

**关键词** 图像配准, 对齐度, 多模态图像

**中图分类号** TN911.73

## 1 引言

图像配准是图像处理的一个基础问题。对同一场景使用相同或不同的传感器(成像设备),在不同条件下(天候、照度、摄像位置和角度等)获取的两个或多个图像一般说会有所不同。简单来说,配准就是对同一场景的不同图像的内容进行空间位置的匹配。多种成像模式产生的图像会表现出不同的分辨率,不同的灰度属性等等差异,它们通常被称为多模态图像。为了对场景进行深入分析,需要把两个或多个多模态图像数据融合起来。实现这些图像的配准,则是最基本的一步。图像配准的应用领域包括:医学图像对比分析,病变定位,大脑或血管造影等;遥感数据分析,农业、地理、海洋、石油、地矿勘探、污染、城市、森林等目标定位、测量和分析;计算机视觉,目标识别、形状重建和运动分析等等。

迄今已报道了多种图像配准方法。例如图像互相关函数(cross-correlation function)<sup>[1]</sup>、相关系数(correlation coefficient)<sup>[2]</sup>、方差求和(sum of squared differences)等方法,基于像素灰度匹配的传统概念,这些方法难于处理多模态图像配准问题。1995年,Viola等人<sup>[3,4]</sup>和Collignon等人<sup>[5]</sup>分别独立地把交互信息(mutual information)引入图像配准领域,为多模态图像配准建议了一种思路,在医学图像处理领域引起了一定的研究兴趣<sup>[6,7]</sup>。该方法要求的计算量大,不适合处理非线性形变和局部形变问题,还有待发展改进。傅氏变换方法已被证明对图像配准十分有用。对图像进行快速傅利叶变换(FFT)后,应用相位相关等技术可以处理旋转、平移和缩放失配的图像配准<sup>[8-10]</sup>。但傅氏变换方法也不能处理非线性变形等问题和不同灰度属性图像的配准。基于特征的图像配准首先要对图像进行预处理,提取可利用的图像特征如角点、线段、闭合或部分边缘曲线、闭合区域重心等,通过寻找这些特征之间的匹配关系达到整幅图像的配准<sup>[11-13]</sup>。然而,例如对于分辨率和灰度属性不同的两幅图像,如何定义它们的特征点的一致性还缺乏成熟的概念。在实用中特征点(控制点)也可人工选取和预先植入。这时,特征点的位置精度受人的主观判断的影响。

文章第2节从进一步分析图像配准的概念入手,并基于此建议一个新的配准准则——对齐度准则,从而将图像配准问题转变为一个参数优化问题。第3节介绍最大化对齐度的多分辨率 Powell 方向族搜索法,第4节展示了一些试验结果,最后是总结与展望。

## 2 新的图像配准准则——对齐度

给定两幅待配准图像  $I_1(x, y)$  和  $I_2(x, y)$ , 我们称其中之一比如  $I_1(x, y)$  为参考图像, 另一

<sup>1</sup> 2002-01-15 收到, 2002-04-15 改回

国家自然科学基金(批准号 60072020)和中国科学院科技创新基金资助

个  $I_2(x, y)$  为待配准图像。在某些图像配准文献中, 图像配准被定义为两个图像之间在坐标位置和灰度级上的双重映射变换<sup>[14]</sup>:

$$I_2(x, y) = g(I_1(f(x, y))) \quad (1)$$

在这里  $f$  是一个二维空间坐标的变换, 而  $g$  是一个一维的灰度变换。

我们认为这个定义是狭隘的。在实际中的多数情况下, 图像配准并不是通过一个或多个变换来达到两幅图像的相等。事实上配准的两幅图像既不需要也常常难于达到等同。

由此我们将图像配准的相关概念明确如下:

图像配准: 给定两幅待配准图像, 对其中一个图像作变换使得变换后的图像与另一个图像的内容在拓扑上相对应并且几何上相对齐 (即建立两幅图像内容的对应关系并使相对应的内容在位置上对齐)。图像配准是以一定的配准准则 (或称匹配准则) 为依据的。图像配准可以称为图像的广义匹配。

按 (1) 式定义的图像配准在本质上是一个狭义的概念, 它所依据的准则可以表示为找  $f$  和  $g$ , 使得最小化下列距离度量:

$$\min_{f, g} \|I_2(x, y) - g(I_1(f(x, y)))\|^2$$

然而, 这种准则不能概括一般的图像配准问题。

在文献中已报道过的多种图像配准方法中, 分别依据了多种配准准则, 如最大化互相关<sup>[1]</sup>或相关系数<sup>[2]</sup>、最小化绝对灰度差、最大化交互信息<sup>[3-5]</sup>、最小化控制点的均方误差<sup>[11-13]</sup>等等。在众多的基于特征的图像配准方法中, 特征控制点的确立过程也体现了该方法在特征点附近的局部图像的匹配准则。比如说图像中的边缘轮廓经编码后的最小匹配误差<sup>[12]</sup>、闭合区域形状的相似度量<sup>[11,13]</sup>、特征的结构相似度量<sup>[15]</sup>等等。我们可以看到, 各个配准准则的侧重点是不同的, 但都力求符合人们某种视觉习惯或要求。

本文提出一个新的匹配度准则: 对齐度。它源自于人眼对两幅图像内容对齐的认识: 当我们判断两幅图像内容对齐的时候, 意味着它们重叠起来交错的区域最小。从微观上来看, 这意味着一幅图像的每个灰度级在像素位置上所对应的另一幅图像的灰度级最稳定, 而这在数学上就体现为方差最小。

假设  $I_1(x, y)$  和  $I_2(x, y)$  大小同为  $M \times N$ , 其灰度直方图分别为  $H_1(n)$  和  $H_2(n)$ , 例如对于有 256 个灰度等级的图像来说,  $0 \leq n \leq 255$ 。对于每一个灰度等级  $n = i$ ,  $H_1(n)$  和  $H_2(n)$  分别代表了灰度值为  $i$  的像素个数, 于是灰度值  $i$  在两幅图像中的出现比率分别为  $p_1(i) = H_1(i)/(M \times N)$  和  $p_2(i) = H_2(i)/(M \times N)$ 。对于  $I_1(x, y)$  的每一个灰度级  $n$ , 现定义  $I_2(x, y)$  相对于  $I_1(x, y)$  灰度值为  $n$  的对应像素集合的灰度均值和方差分别为  $\bar{E}_{1,2}(n)$  和  $\sigma_{1,2}^2(n)$ 。其表达式分别为

$$\bar{E}_{1,2}(n) = \frac{1}{H_1(n)} \sum_{I_1(x,y)=n} I_2(x, y) \quad (2)$$

$$\sigma_{1,2}^2(n) = \frac{1}{H_1(n)} \sum_{I_1(x,y)=n} (I_2(x, y) - \bar{E}_{1,2}(n))^2 \quad (3)$$

同理, 对于  $I_2(x, y)$  的每一个灰度级  $n$ , 定义  $I_1(x, y)$  对于  $I_2(x, y)$  灰度值为  $n$  的对应像

素的灰度均值和方差分别为  $\bar{E}_{2,1}(n)$  和  $\sigma_{2,1}^2(n)$ 。其表达式分别为

$$\bar{E}_{2,1}(n) = \frac{1}{H_2(n)} \sum_{I_2(x,y)=n} I_1(x,y) \quad (4)$$

$$\sigma_{2,1}^2(n) = \frac{1}{H_2(n)} \sum_{I_2(x,y)=n} (I_1(x,y) - \bar{E}_{2,1}(n))^2 \quad (5)$$

对  $\sigma_{1,2}^2(n)$  以灰度值  $n$  的出现比率  $p_1(n)$  进行加权平均得到基于  $I_1(x,y)$  的期望方差定义为  $\bar{\sigma}_{1,2}^2$  :

$$\bar{\sigma}_{1,2}^2 = \sum_n p_1(n) \sigma_{1,2}^2(n) \quad (6)$$

同理, 对  $\sigma_{2,1}^2(n)$  以  $p_2(n)$  进行加权平均得到基于  $I_2(x,y)$  的期望方差  $\bar{\sigma}_{2,1}^2$  :

$$\bar{\sigma}_{2,1}^2 = \sum_n p_2(n) \sigma_{2,1}^2(n) \quad (7)$$

由此我们定义两幅图像的交互方差 CI 为

$$CI[I_1, I_2] = \bar{\sigma}_{1,2}^2 / \sigma_2^2 + \bar{\sigma}_{2,1}^2 / \sigma_1^2 \quad (8)$$

其中,  $\sigma_1^2$  和  $\sigma_2^2$  分别为图像  $I_1(x,y)$  和  $I_2(x,y)$  方差, 定义为

$$\sigma_1^2 = \frac{1}{MN} \sum_{(x,y)} (I_1(x,y) - \mu_1)^2, \quad \mu_1 = \frac{1}{MN} \sum_{(x,y)} I_1(x,y) \quad (9)$$

$\sigma_2^2$  类似地定义。我们说, 这个交互方差 CI 就反映了两幅图像灰度相互对应的稳定程度。以  $\sigma_1^2$  和  $\sigma_2^2$  作分母可以去除图像整体方差的大小产生的影响。进一步的分析可以证明交互方差满足以下性质:

(1) 对称性  $CI[I_1, I_2] = CI[I_2, I_1]$

(2)  $CI[I_1, I_2] \geq 0$  在无噪声情况下, 如果  $I_1(x,y)$  和  $I_2(x,y)$  的棱边能达到理想对齐, 则  $CI[I_1, I_2]$  达到最小值 0, 而与  $I_1(x,y)$  和  $I_2(x,y)$  的灰度属性无关。

(3)  $CI[I_1, I_2] \leq 2$  容易证明:  $\bar{\sigma}_{1,2}^2 \leq \sigma_2^2$  和  $\bar{\sigma}_{2,1}^2 \leq \sigma_1^2$ 。在两幅图像完全不相关的情况下, 它们的交互方差将接近 2。

可以看出, 两幅图像内容越对齐, 它们的交互方差越小, 而这一关系的建立并不需要两幅图像的灰度之间有线性的相关, 也不受两幅图像灰度属性差异的影响。为了实验中描述方便起见, 我们将交互方差 CI 的倒数称为两幅图像的对齐度 (AM, Alignment Metric) :

$$AM[I_1, I_2] = \frac{1}{\bar{\sigma}_{1,2}^2 / \sigma_2^2 + \bar{\sigma}_{2,1}^2 / \sigma_1^2} = \frac{\sigma_1^2 \sigma_2^2}{\sigma_{1,2}^2 \sigma_1^2 + \sigma_{2,1}^2 \sigma_2^2} \quad (10)$$

并以此作为新的图像匹配准则, 提出基于最大化 AM 的图像配准方法。

一维数字信号可以看作是二维图像的一种特殊形式, 即其中行或列的尺寸大小为 1。因此一维数字信号的对齐度可以类似方法求得。同理可以很容易地将 AM 扩展到三维以及更多维, 定义多维图像的 AM。

图 1 给出了两列数字信号, 它们的变动边缘位置相互对应, 但若干电平值有可见的不同, 并含有噪声。图 2 给出了图 1 中两列数字信号相互移位产生的对齐度序列, 可以看到在对齐位置有明显的极值。

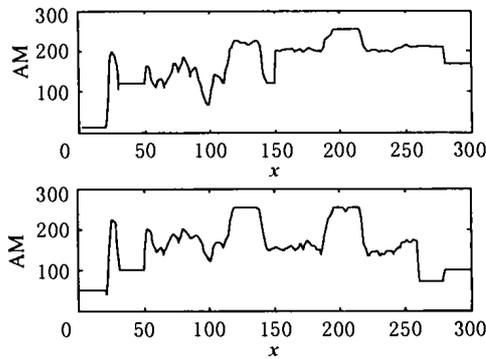


图1 两列可配准的一维数字信号

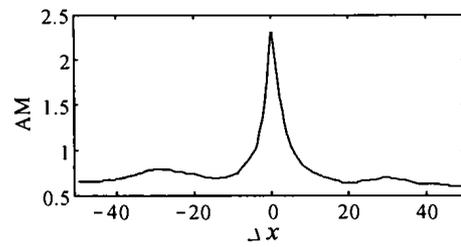


图2 图1中两列数字信号的对齐度

最大化 AM 作为一个新的图像配准方法具有以下主要特点: (1) 具有简单的计算复杂度, 不需要建立概率模型, 因此计算速度快; (2) 不需要配准图像的灰度之间有相关性, 可以配准多模态图像, 包括有很大灰度属性差异和直方图差异的图像; (3) 对齐度可以有效适用于一维数字信号和多维图像的配准; (4) 本方法不敏感于均匀分布的白噪声, 尽管噪声会对 AM 的峰值的尖锐程度造成一些负面影响, 本文第 4 节对此会有进一步的讨论。

### 3 最大化对齐度

我们说, 图像配准的目的是找到一个变换  $T$ , 使得对  $I_1(x, y)$  作变换后的图像  $I_T(x, y) = T\{I_1(x, y)\}$  与  $I_2(x, y)$  的对齐度最大, 即最大化  $AM[T\{I_1\}, I_2]$ 。这里变换  $T$  是一个二维空间域坐标的几何变换, 可以表示为

$$(x', y') = T_{\xi}(x, y) \quad (11)$$

其中  $\xi$  为变换  $T$  的参数集。变换  $T$  既可以是线性的, 也可以是非线性的。它通常可以用多种变换模型来假定, 常用的变换模型有: 3 参数旋转平移 (刚体变换)、4 参数旋转平移缩放、6 参数仿射变换以及多项式变换等等。本文考虑刚体变换的情况, 即变换由旋转和平移构成, 包含 3 个参数:  $\xi = (\Delta x, \Delta y, \Delta \theta)$ , 其中  $\Delta x$  和  $\Delta y$  反映了平移量,  $\Delta \theta$  反映旋转量。我们可以把对齐度 (AM) 写成参数集  $\xi$  的函数, 表示为

$$AM(\xi) = AM[T_{\xi}\{I_1\}, I_2] \quad (12)$$

配准问题转化为寻找最优参数集  $\xi^*$  使得  $AM(\xi^*)$  达到最大, 表达为

$$\xi^* = \arg \max_{\xi} AM[T_{\xi}\{I_1\}, I_2] \quad (13)$$

如同大多数的多参数优化问题一样,  $\xi^*$  的确定通过迭代搜索来完成。有多种极值搜索方法可以应用。本文应用 Powell 方向族搜索法<sup>[16]</sup>来搜索 AM 的全局最大值。Powell 的方法通过以一定顺序在各个一维方向上的迭代搜索来完成多维函数的最大化, 对此方法更详细的介绍可参见文献 [16]。比如对本文的三维变换  $T_{(\Delta x, \Delta y, \Delta \theta)}$  来说, 本文以  $\Delta x$ - $\Delta y$ - $\Delta \theta$  的顺序搜索, 先在  $\Delta x$  方向搜索到极值, 固定这个  $\Delta x$  值, 然后在  $\Delta y$  方向搜索到极值后固定  $\Delta y$ , 然后是  $\Delta \theta$  方向, 之后再回到  $\Delta x$  方向迭代搜索, 直到收敛到所搜索的参数集  $\xi^*$  或者超过了规定的最大搜索次数。实验证明, 如果有接近最优参数集  $\xi^*$  的初值  $\xi_0$  作为 Powell 方法的搜索起始点的话, 搜索收敛速度将极大提高, 本文用多分辨率方法来解决初值问题。

数字图像的多分辨率分析是近些年在数字图像处理领域广泛使用的显著提高处理速度和精度的方法, 本文也将其应用到图像配准上。其基本思路是首先在最低的图像分辨率下获得问题的初始解。这时由于图像小, 计算量很小, 然后在更高的分辨率下实施处理, 将前一步低分辨率下的处理结果内插获得高分辨率下的初始图像。继续这个过程达到最高分辨率图像。本文在最低分辨率图像上经全搜索找到全局极值对应的参数集, 再将此参数集映射到高分辨率层次作为搜索的起始初值, 使用 Powell 方法搜索极值, 再向更高分辨率继续。实验证明它可以有效和稳定地解决全局最优参数集搜索问题。

### 4 实验结果与分析

本文分别用遥感图像和医学图像进行配准实验。图 3 是裁剪于两个不同时间的 TM 卫星图像分别为 09/09/90 (TM905A) 和 07/18/94(TM945A)。它们之间有平移失配, 灰度属性也有可见的差别。图 4(a) 给出了不同  $\Delta x$  和  $\Delta y$  对应的 AM 值, 可以看到配准位置有明显的极值, 而且 AM 曲面是光滑的。图 4(b) 是极值点对应的配准结果。

图 5(a) 和图 5(b) 给出了两幅有旋转失配的多模态图像, 分别为 MR 图像和 CT 图像, 图 5(c) 是对应不同旋转角度的对齐度。图 6 给出了 CT 图像添加 10dB 噪声后的配准结果。实验表明噪声没有影响配准结果, 只影响了 AM 曲线在极值点的锐度。图 7 展示了配准结果的边缘对齐程度。

图 8 给出了有旋转平移失配的两幅 TM 图像, 分别取自 07/18/94(TM945A) 和 09/09/90 (TM905A)。使用多分辨率的 Powell 方向族法搜索得到变换参数集为  $\Delta x = 84$ ,  $\Delta y = -36$ ,  $\Delta\theta = -10.5$ , 图 9 给出了相应参数集的配准结果。

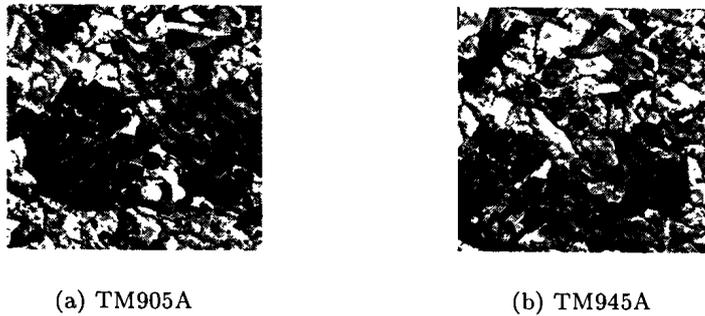


图 3 TM 图像

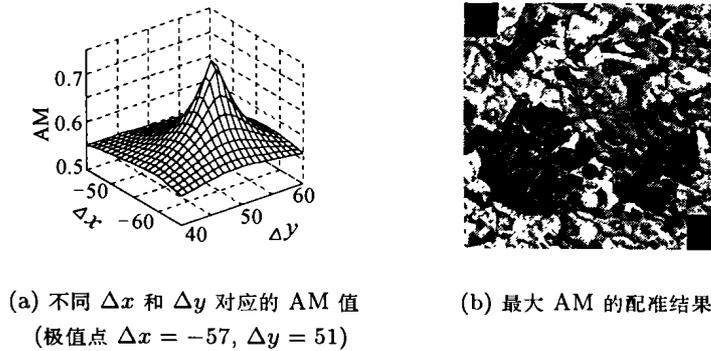


图 4



## 6 总结与展望

本文从图像配准的定义出发, 提出了一种新的图像配准则: AM。它能够正确地反映图像之间的配准程度而不需要图像的灰度之间有线性相关, 因而可以基于最大化 AM 来配准有极大对比度差异的图像和多模态图像。它还适用于一维数字信号或多维图像的配准。本文使用多分辨率方法和 Powell 方向族搜索法来最大化 AM。多种实验结果证实了该准则的有效性和优点。作者认为值得对此进行更深入的研究和拓宽其应用领域。

## 参 考 文 献

- [1] A. 罗申菲尔特, A. C. 卡克著, 余英林等译, 数字图像处理, 北京, 人民邮电出版社, 1982, 第 1, 2 章.
- [2] M. Svedlow, C. D. McGillem, P. E. Anuta, Experimental examination of similarity measures and preprocessing methods used for image registration, Symposium on Machine Processing of Remotely Sensed Data, Purdue University, Indiana, June 1976, 4A-9.
- [3] P. A. Viola, W. M. Wells III, Alignment by maximization of mutual information, Proc. 5th Int. Conf. Computer Vision, Boston, MA, June, 1995, 16-23.
- [4] P. A. Viola, W. M. Wells III. Alignment by maximization of mutual information, International Journal of Computer Vision, 1997, 24(2), 137-154.
- [5] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, G. Marchal, Automated multi-modality image registration based on information theory, Proc. of the Information Processing in Medical Imaging Conference, Norwell, MA: Kluwer, 1995, 263-274.
- [6] Jsiem P. W. Pluim, J. B. Antoine, Max A. Viergever, Image registration by maximization of combined mutual information and gradient information, IEEE Trans. on Medical Imaging, 2000, 19(8), 809-814.
- [7] Philippe Thevenaz, Michael Unser, Optimization of mutual information for multiresolution image registration, IEEE Trans. on Image Processing, 2000, 9(12), 2083-2098.
- [8] S. Alliney, Spatial registration of multispectral and multitemporal digital imagery using fast-Fourier transform techniques, IEEE Trans. on Pattern Analysis and Machine Intelligence, 1993, 15(5), 499-504.
- [9] S. Alliney, C. Morandi, Digital image registration using projections, IEEE Trans. on Pattern Analysis and Machine Intelligence, 1986, 8(2), 222-233.
- [10] B. S. Reddy, B.N. Chatterji, An FFT-based technique for translation, rotation, and scale-invariant image registration, IEEE Trans. on Image Processing, 1996, 5(8), 1266-1271.
- [11] J. Flusser, T. Suk, A moment-based approach to registration of images with affine geometric distortion, IEEE Trans. on Geoscience and Remote Sensing, 1994, 32(2), 382-387.
- [12] H. Li, B.S. Manjunath, S. K. Mitra, A contour-based approach to multisensor image registration, IEEE Trans. on Image Processing, 1995, 4(3), 320-334.
- [13] A. Goshtasby, G. C. Stockman, C. V. Page, A region-based approach to digital image registration with subpixel accuracy, IEEE Trans. on Geoscience and Remote Sensing, 1986, GE-24(3), 390-399.
- [14] L. G. Brown, A survey of image registration techniques, ACM Computing Surveys, 1992, 24(4), 325-376.
- [15] A. D. Ventura, A. Rampini, R. Schettini, Image registration by recognition of corresponding structures, IEEE Trans. on Geoscience and Remote Sensing, 1990, 28(3), 305-314.
- [16] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, Cambridge, U.K., 1988, 309-317.

## AUTOMATIC REGISTRATION OF MULTI-MODAL IMAGES BASED ON THE MAXIMIZATION OF THE ALIGNMENT METRIC

Wang Dongfeng    Zou Mouyan

*(Institute of Electronics, Chinese Academy of Sciences, Beijing 100080, China)*

**Abstract** This paper discusses the generalized definition of image registration and proposes a new measure of image registration—the Alignment Metric(AM). This metric can effectively describe the degree of match of two images with different gray level properties. As a result, the problem of image registration is turned into the maximization of the AM. The new criterion can be used to register multi-modal images including the images with different gray level properties. The Powell's direction set method with multi-resolution is used to search for a global maximum of the AM. The images can then be registered automatically without the need of the previously located control points or landmark points. Experiments using multi-modal medical and remote sensing images demonstrate the effectiveness of the method.

**Key words** Image registration, Alignment Metric(AM), Multi-modal images

王东峰: 男, 1974年生, 助理研究员, 研究方向为图像分析和处理, 数字信号处理.

邹谋炎: 男, 1941年生, 研究员, 博士生导师, 研究方向为信号和图像处理以及宽带无线通信.