Multimodal Data Fusion Based on Mutual Information

基于互信息的多模态数据融合

转述人：常家鑫

Bramon R, Boada I, Bardera A, et al. Multimodal data fusion based on mutual information[J]. IEEE Transactions on Visualization and Computer Graphics, 2011, 18(9): 1574-1587.

**摘要**

多模态可视化的目的是融合不同的数据集，使产生的组合为用户提供更多的信息。为了实现这一目标，我们提出了一种新的信息理论方法，从两个体积数据集中自动选择信息量最大的三维像素。我们的融合标准是：基于两个输入数据集之间创建的信息通道，允许我们量化与每个强度值相关的信息。这个具体的信息是利用三种不同的分解通道相互信息的方式获得的。此外，基于融合数据集的信息含量的评估标准可以通过加权每个数据集对最终结果的贡献，用来分析和修改初始选择的三维像素。本文所提出的方法已经集成在一个总体框架中，该框架允许探索体积数据模型和融合数据集的一些参数的交互变化。所提出的方法已经在不同的医学数据集上进行了评估，并取得了非常可喜的成果。

**引言**

目前的医学影像设备能够获取身体任何器官的详细信息，这使得医学影像和三维容积可视化成为病人诊断和治疗的标准。更重要的是，放射影像能够代表补充信息。鉴于大多数疾病的放射学模式的异质性和复杂性，自动结合计算机断层扫描（CT）、磁共振成像（MRI）或正电子发射断层扫描（PET）的最相关信息的多模态成像新方法，可以发现并检测在单独考虑某种模态情形下可能会被错过的信息。在这种情况下，多模态可视化已成为研究的重要焦点。

多模态可视化的目的是将来自不同体积数据集的最相关信息结合为单一的可视化，并提供尽可能多的信息。为了达到这个目标，必须进行两个主要步骤。首先，需要一个配准步骤，将输入的数据集排列在一个共同的空间中。配准已经被广泛研究，并提出了许多不同的算法。在多模态配准中，最流行的算法是基于两幅图像之间相互信息的最大化。其次，还要进行融合步骤，将同一空间位置上的代表值进行混合，以得到最终的图像。

融合过程具有挑战性，因为它要求将每对配准三维像素中包含的数据还原成单一的视觉表示。已经提出了不同的替代方法来执行信息合并并生成融合图像。最简单的方法是棋盘显示，交替显示原始输入模型的像素，基于颜色通道的融合，将输入图像分配到不同的颜色通道，然后执行渲染，以及数据分割（表示其中一个模型的信息）与颜色编码（映射另一个模型的数据）的组合。这些技术已经被证明在临床环境中是有用的，不过，在大多数情况下，重建结构之间关系的任务是留给观察者的。

在医疗环境中，最关键的环节不是 不仅要找到合适的融合方案，还要找到一个有用的视觉化方案。这种可视化必须使我们能够建立输入模型之间的视觉对应关系，同时保持不同来源的特征。 因此，可视化必须在不同输入模型的互补方面提供尽可能多的视觉线索。另一方面，视觉化必须是可重现性，因为不同的用户可以获得相同的结果。

在本文中，我们提出了一种新的多模态可视化方法，在提取输入数据集最相关信息的基础上自动融合数据。首先，我们在两个配准数据集之间设置一个信息通道。然后，我们使用不同的基于信息的算法来选择输入数据集中信息量最大的三维像素。这个信息量是由三个不同的与强度值相关的具体信息度量计算出来的。这些度量对应于通道的相互信息的不同分解。一旦选择了一个三维像素，就会使用与其对应的数据集相关联的原始传递函数将其可视化。最后，基于融合数据集的信息内容的标准允许我们通过调节每个数据集对最终结果的贡献来修改初始选择的三维像素。所提出的方法被集成到一个总体框架中，允许我们探索体积数据，并交互式地修改融合中涉及的一些参数。所提出的框架已经在不同的临床数据集上进行了评估。

**本方法的动机和概述**

数据集融合的主要目标是将不同的源数据集紧凑地表示成一个包含每个源数据集最重要特征的单一图像。融合后的数据集也应该比任何一个输入数据集包含更多的信息。在本文中，我们提出了两种多模态数据集的融合策略，对于每一对对应的三维像素，自动选择信息量最大的强度值，这样融合后的数据集将只由来自源图像的值构成（即不创建新的值）。

现现在出现了一些问题。我们能否找到一种自动融合策略，从源图像中收集最大的信息？我们能不能通过观察另一幅图像来获取信息？如果我们在两个配准的多模态数据集之间建立一个通信通道，从信息理论中可以知道，相互信息（MI）量化了它们之间传输或共享的信息量。在融合的背景下，需要强调的是，MI表达的是通过观察对方获得的图像信息。另一个相关的事实是，MI可以以不同的方式分解，以获得与每个强度值相关的特定信息。在本文中，三种特定信息的测量方法被应用于新的融合标准的定义，从不同的角度选择信息量最大的三维像素来产生融合数据集。这些测量方法已经被引入神经系统领域，用于研究、刺激和反应相关的信息。融合模型的可视化是由每个输入模型的预定义传递函数获得的。最后，我们提出了一个评估标准来研究和调控融合数据集的信息含量（熵）。因此，我们提出了从两个不同的数据集获得融合数据集的五个步骤（见图1）。

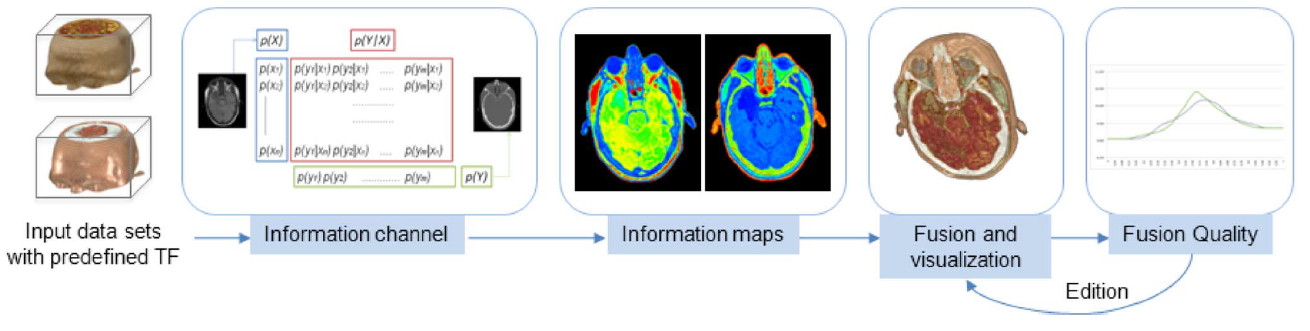


图1 多模数据融合管道概述

1. 通信渠道。在两个配准的多模态数据集之间定义了一个通信通道。这个通道允许我们考虑数据集之间的相互信息。

2. 具体信息。对于每个数据集，计算出与每个强度值相关的信息。这种信息也称为具体信息。分别从三种不同的相互信息分解中得出的三种信息度量。对于每个源数据集，创建一个信息图，代表与每个强度值相关的具体信息。

3. 融合标准。融合图像是从源数据集中信息量最大的三维像素中获得的。也就是说，对于每一对匹配的三维像素，选择信息量最大的一个（即具有最高的特定信息）作为融合数据集。

4. 视觉化。多卷射线投射用于渲染融合模型。颜色和不透明度从为每个输入模型定义的原始转移函数中获得。

5. 评估标准。提出了一个基于融合图像熵的评估标准，以研究和调节每个数据集对最终结果的贡献。

**结果**

在本节中，我们描述了开发的平台，以进行不同的实验，已进行评估所提出的方法。所获得的结果已由一组专家（巴塞罗那医院诊所的一名解剖学家，以及赫罗纳Josep Trueta医院的两名神经放射学家和一名神经外科医生）进行了分析。

**1. 多模态可视化平台**

我们开发了一个多模态可视化平台，将已经提出的不同方法整合在一起。使用基于CPU的射线投射来渲染输入模型和信息地图。融合模型使用基于CPU的多模态射线铸造渲染，在积累过程中执行混合。

**2.测试数据集**

为评估所提出的方法，使用两套不同的大脑图像。第一个包含已从赫罗纳的Josep Trueta医院的三个不同的病人那里获得的CT和T1图像。按照机构审查委员会对人类课题研究的指导方针，所有受试者均已知情同意。这些模型的分辨率是512\*512\*26三维像素的CT和288\*288\*20三维像素的T1。第二个测试数据集对应于从回顾性图像配准评估项目数据库中标记为训练的患者。从这个患者中我们选择CT、T1、T2和PET图像。这些模型的分辨率为：CT为512\*512\*26个三维像素，MRI序列（T1和T2）为256\*256\*26个三维像素，PET为128\*128\*15三维像素。

根据这些数据集，我们评估了CT-T1、T1-T2和CT-PET融合。在预处理步骤中，测试模型被配准并重新采样到图像对的最低分辨率，并且去除背景。然后，分别对所提出的对称和非对称方法进行了应用。专家们利用开发的平台对所得结果进行了分析。在评估过程中，专家们并不知道所采用的融合策略。

在评估之前，专家们为每一对模型确定了最相关的信息，同时也确定了他们将对最终的融合进行评估的内容。这些信息可以总结为以下几点。CT能探测到致密的结构，如骨骼，给出物体的大致形状，但对软组织的细节却很少。T1图像用于描述软组织的形态是丰富的细节。在CT-T1融合中，医生希望从CT上看到致密的结构，从T1上看到软组织，在T1-T2对中，T1图像提供了解剖信息，给出了组织的大致形状，T2图像捕捉了组织中水分的信息，很适合检测含水量增加。在T1-T2融合中，组织形状最好从T1采集，水分异常从T2采集。在CT-PET对中，PET提供了代谢活动模式的信息，而CT提供了高质量的空间背景信息。在CT-PET融合中，医生希望从PET中看到实质，从CT中看到骨和其他结构。另一方面，为了评价最终的融合结果，专家们从边界和内容上对信息进行了描述。在结构（皮肤、肌肉、骨骼）之间最好能确定边界，在内容方面，最好能确定实质的细节，如裂隙和沟槽、实质的内部结构。

**3.对I1、I2和I3的评价**

第一个实验是为了说明信息度量I1、I2和I3的行为。将不同的融合策略应用于测试数据集。图2显示了CT-T1(列(a-b))、T1-T2(列(c-d))和CT-PET(列(e-f))的输入模型(行(i))，以及I1、I2和I3的信息(行(ii-iv))。

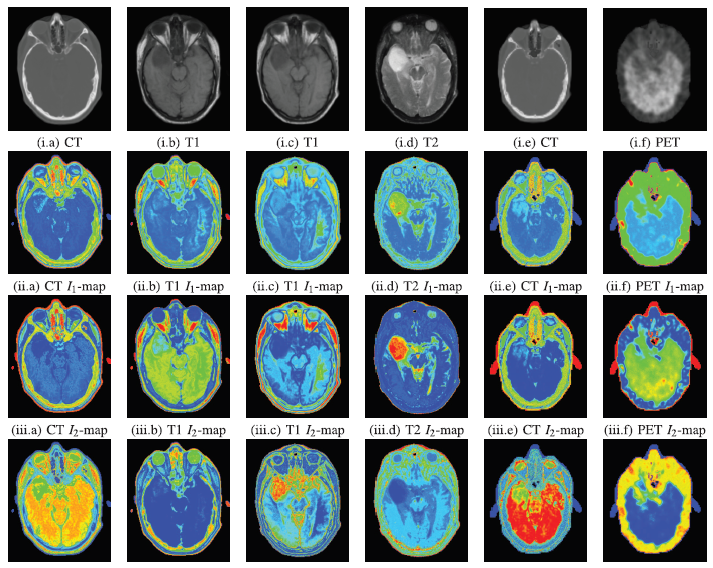


图2 CT-T1、T1-T2、和CT-PET信息图

图2(ii)表示了度量I1的行为。这个度量代表给定值x相关强度的分布与平均分布的不同。在对CT-T1(图2(ii.a-ii.b))中，I1的最低值对应于实质，而最高值则是在眼外肌和副鼻窦中实现的。我们观察到，在CT-T1的情况下，软组织，骨结构和实质之间的边界更好地定义。这种效果是由于CT和T1图像中的信息比其他对模中的信息互补性更强。在T1- T2这对图像中（图2(ii.c-ii.d)），测量值I1相当低。观察T2中的病理病变达到了很高的数值，说明它是一个信息量很大的部分。在对CT-PET(图2(ii.e-ii.f)中，由于CT图像提供解剖信息，PET图像提供功能信息，所以两个I1图的特征是非常不同的。

衡量标准I2代表可预测性，即当观察到x时，Y不确定如何变化。专家们强调了划分结构的能力。对于CT-T1这一对(图2(iii.a-iii.b))，观察CT-脉络体的I2值较低，而T1-实质体的I2值较高。因此，观察CT-脉络质产生的不确定度比T1小，而观察T1-脉络质产生的不确定度比CT高。这主要是由于T1比CT更详细地捕捉了脑组织的信息。因此，融合方法将选择T1像素导致的图像，其中皮肤，骨结构，肌肉和沟槽是非常好的划分。对于T1-T2这对图像（图2(iii.c-iii.d)），我们观察到，在T2图像的病理区域实现了I2的最高值，这比使用I1和I3更好的分辨。还需注意病变（蛛网膜囊肿）的轮廓如何在T2图像中清晰显示（图2（iii.d））。对于一对CT-PET（图2（iii.e-iii.f），测量I2的表现与I1相似，尽管I2更好地突出了结构的边界，这是解释图像数据时的一个理想特征。

最后，图2(iv)给出了度量I3的行为。这个度量表示当观察到与x相关的强度y时，关于x的不确定性如何变化。这种方法的主要局限性是边界的不连续性的出现，造成模型探索的混乱。

所有专家都一致认为，最好的信息图是通过测量I2实现的(图2(iii))。

**4. 对称和非对称融合评估**

在接下来的实验中，我们使用第2节中描述的数据集和不同的k值来评估对称和非对称融合方法的行为。展示使用不同融合策略得到的融合数据集。

四个CT-T1测试数据集的H(z)行为如图3所示，其中(a-c)分别是对称方法I1-融合、I2-融合和I3-融合对应的图，(d-e)是非对称方法I2I3-融合的图。在所有的图中，X轴代表值k(范围从0到1)，Y轴代表评估度量H(z)的值。请注意，对于所有的方法来说，k接近于0或1时，图中的极端值具有平面形状。在这种情况下，融合模型的信息完全来自CT或T1，因此熵H(z)的值与相应输入数据集的熵相吻合。从对称方法的结果(图6a、6b和6c)中，我们观察到，在I1和I3的情况下，当k接近0.5时，H(z)达到最大值，而对于I2，在很宽的数值范围内，熵接近最大值。这种行为赋予了I2-融合的高鲁棒性。对于非对称方法I2-I3-融合(图6d和6e)，熵H(z)的行为与I2-融合类似，也表现出较高的鲁棒性。

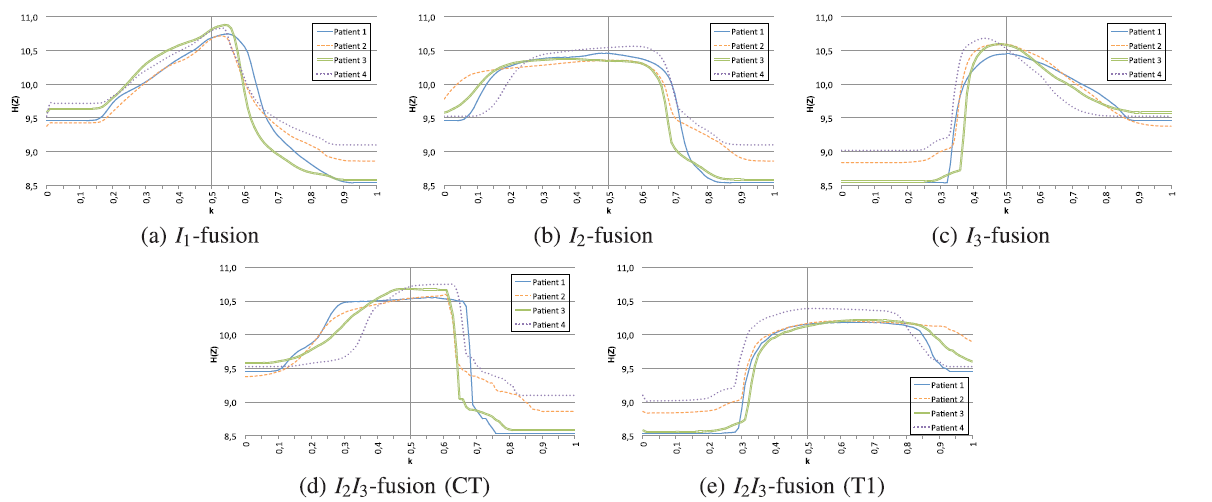


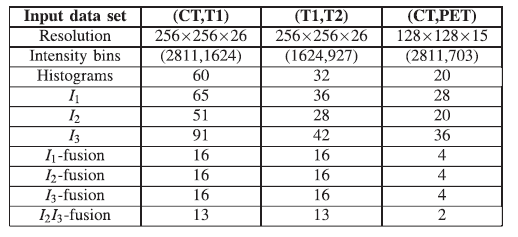
图3 所提出融合策略下，H(z)在不同CT-T1 pairs下关于k值的表现

**5.计算时间**

所有的测试都是在MacBook Pro笔记本电脑系统上进行的，拥有运行在266GHz的英特尔 酷睿i7 CPU、4GB内存和nVidia GeForce GT 330M 512 MB显卡。

我们的性能测试结果如表1所示。对于每一对模型，我们报告了融合模型的分辨率，每个输入模型的强度宾格数，以及计算直方图、措施I1、I2和I3以及考虑不同策略的融合模型所需的时间（毫秒）。观察到，即使CT、T1和T2具有相同的分辨率，CT-T1融合的直方图和度量衡的计算时间比T1-T2融合的时间要高。这是由于这些计算不是取决于其分辨率。此外，我们可以看到，I2I3融合方法比对称方法更快，因为它不需要在每个像素询问两个模型的信息。考虑到不同的k值，可以得到融合模型，但修改参数k不需要重新计算。需要注意的是，渲染前所需的计算速度非常快，因此，所提出的技术允许交互式帧率。在用户版的情况下，交互式地获得融合结果，可以根据用户的操作进行即时的视觉反馈。

表1 计算直方图



**6.多卷处理**

最后，我们介绍一些自动获得的三维多模态可视化。为了说明所提出的方法的良好性能，我们分别展示了使用棋盘方法渲染的CT-T1对（图4a），输入值的加权平均（图4b），以及用I2-融合和I2I3(CT)-融合方法获得的融合模型的渲染（图4c和4d）。观察到，所提出的融合方法比经典方法产生的图像对比度更高。

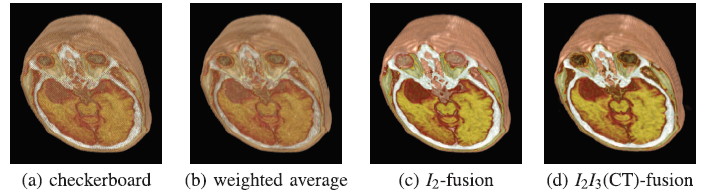


图4 CT-T1融合模型的三维可视化

**结论**

在本文中，我们提出了一个新的信息理论框架来融合两个多模态图像。提出的融合标准是基于在两个配准数据集之间设置一个信息通道。从这个通道中，我们分析了与每个强度值相关的三个特定信息测量的行为，我们提出了两种算法策略，选择信息量最大的体素来获得融合图像。我们还引入了一个基于融合图像熵的评估标准，通过调节每个图像对最终结果的贡献来修改初始选择的像素。所提出的方法已经在不同的医疗数据集上进行了评估，所得到的结果已经由一组医生进行了分析，认为所提出的融合方法对规划放疗、治疗监测和规划脑部手术有潜在的作用。

作为未来工作的基础，我们将研究如何将这种方法推广到两个以上数据集的融合。这种扩展需要对两个以上变量的互信息分解进行详细分析。我们还将分析如何在融合框架中考虑数据集的空间相干性和梯度信息。我们还将对不同图像模式和解剖结构的特定信息测量的行为和可用性进行更详尽的分析。我们计划在一个大型患者队列中测试和验证这种新方法，以确定其临床效果。最后，我们将探索如何将信息图谱用于定义信息传递函数。