# 一种基于本体概念语义距离的服务相似度度量方法

葛继科 邱玉辉
(西南大学计算机与信息科学学院 重庆 400715) (西南大学语义网格实验室 重庆 400715)

**摘 要** 随着语义Web服务及语义网格服务应用的不断深入,对服务资源的需求日益增长,服务匹配在服务发现和服务组合研究中的地位也日渐重要。在服务使用OWL-S描述的前提下,服务匹配通常认为是本体概念的匹配,概念匹配的目的是发现概念间的语义相似度。概念的语义相似度不但与概念间的距离有关系,而且还受概念在本体中层次深度的影响。综合考虑这两个因素,提出了一种基于语义距离的概念相似度度量方法,给出了语义距离的定义,明确了语义距离与语义相似度的关系。最后,通过与其他方法的实验比较,验证了该方法的有效性。

**关键词** 服务匹配,语义相似度,本体,语义距离
**中图法分类号** TP301 **文献标识码** A

## Service Similarity Measure Based on Semantic Distance of Ontology Concepts

GE Ji-ke QIU Yu-hui
(Faculty of Computer and Information Science, Southwest University, Chongqing 400715, China)
(Semantic Grid Research Group, Southwest University, Chongqing 400715, China)

**Abstract** With the application of semantic Web service and semantic Grid service, the requests of service resource are on the fly, and service matching is more and more important in the domain of service discovery and service composition. Under the precondition of service described by OWL-S, service matching are generally considered as the matching of ontology concepts, and the purpose of concepts matching is that finding semantic similarity between concepts. The semantic similarity between concepts is influenced not only by the distance between concepts, but also by position of concept in the ontology hierarchy architecture. In this paper, considering the above two factors, we proposed a concepts similarity measure based on the semantic distance, defined the semantic distance, and stated the relationship between the semantic distance and semantic similarity. At last, we also provided an experimental comparison of our measure against traditional similarity measures, and proved empirically the benefits of our approach.

**Keywords** Service matching, Semantic similarity, Ontology, Semantic distance

## 1 引言

近年来,语义Web及语义网格环境中的服务匹配成为一个新的研究热点,匹配操作对两个实体从语义上产生一个映射[1]。在Web服务使用Web Ontology Language for Service (OWL-S)[2]描述的前提下,服务匹配通常被认为是基于本体概念的语义相似度计算。

对于概念的语义相似度计算,国外许多研究者利用了语义词典WordNet中的同义词集组成的树状层次体系结构[3],总体上分为两类方法:一类是考虑两个概念共享信息的程度,基于信息理论定义相似度计算方法;另一类是采用先计算两个概念在树中的语义距离,然后转化为语义相似度的办法。

SUMO (Suggested Upper Merged Ontology)[4] 是由IEEE标准上层知识本体工作小组所建置的,其目的是发展标准的上层知识本体,这将促进数据互通性、信息搜寻和检索、自动推理和自然语言处理。在SUMO中,所有的概念根据语义被组织在一棵层次树中。所以,要度量其中任意两个概念的语义相似度,可以先计算这两个概念在层次树中的语义距离,然后再将语义距离转换为语义相似度。这里的语义距离是指两个概念的相近程度,一般说来,两个概念间的语义距离越小,它们的语义越相近,反之越远。在SUMO层次树中,自顶向下,概念的分类是由大到小,大类间的概念相似度一般要小于小类间的。因此,在同等语义距离的情况下,处于层次树中离根较远的概念间的相似度要比离根近的概念间相似度大。由此可见,概念在树中所处的深度在语义相似度度量时也是一个需要考虑的因素。本文设计了一种考虑上述因素的语义距离计算方法,以求能够获取有价值的语义相似度。

## 2 相关理论

### 2.1 本体的形式化描述

到稿日期:2008-09-21 返修日期:2008-11-11 本文受973国家重大基础研究项目(2003CB317008),西南大学研究生科技创新基金(2006011)资助。

葛继科(1977—),男,博士生,主要研究方向为语义网格、服务发现等,E-mail:gjkid@swu.edu.cn;邱玉辉(1938—),男,教授,博士生导师,主要研究方向为语义网格、服务发现、人工智能等。

本体是共享概念化的明确的形式化规范[5]。本体主要描述概念间的层次结构，上层概念的语义相对于底层概念更为抽象，共享的程度高；而底层概念较为具体，更贴近具体的应用，概念之间是泛化和特例的关系。在许多应用环境中，概念之间的关系通过语义进行关联，本体可以作为表达这种语义关联的框架。

**定义 1**(本体) 本体 $O$ 是一个 5 元组 $O:=(V,F,C,H,Root)$，其中 $V$ 是一组词汇集；$C$ 是一组概念；$F$ 是一个参照函数，将一个词汇集 $\{V_i\}\subset V$ 映射到一个概念集。通常来说，多个词汇可以映射到一个概念，一个词汇也可以映射到多个概念；$H$ 是层次关系 $H\subseteq C\times C$，$H(c_1,c_2)$ 代表 $c_1$ 是 $c_2$ 的子概念，$H$ 是有向的、无环的、传递的、自反的；$Root$ 是一个根概念，$\forall c\in C, H(c,Root)$ 成立，在一个本体中，有且只有一个 $Root$。如未明确指出，一般认为 $H(c_1,c_2)$ 关系中，$c_1\neq c_2$。图 1 描述了一个简单本体的示意图。

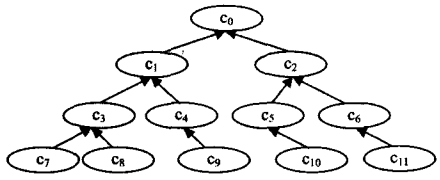


图 1 一个简单的本体示意图

### 2.2 相似度计算方法

在服务匹配的相关研究领域，一些有效的相似度匹配方法已被提出，主要分为如下几类。

#### 2.2.1 基于字符串的方法

该类方法使用构词法相似性来寻找概念间的相似度，主要反映了两个概念在语言学上的相近程度[6]。如，“book”和“textbook”具有较高的相似度。但当两个具有相同内涵的概念具有不同的语言形式时，该方法将无法正确度量它们的相似性。如，该类方法很难发现“book”和“volume”之间的相似性。

#### 2.2.2 基于同义词词典的方法

该方法根据同义词词典将所有的概念组织在树状的层次结构中，任意两个节点之间有且只有一条通路，这条通路的长度就作为这两个概念间语义距离的一种度量[7,8]。这种方法简单有效，但其结果受人的主观意识影响较大。而且概念的层次结构多数并不是树状的，同时概念节点之间可能不止一条通路。

#### 2.2.3 基于特征匹配的方法

该类方法依据概念或对象的特征来判断语义相似度。如文献[9]从语言学角度研究了概念相似度，提出了相似度不仅由两个概念的相同属性决定，而且由它们的不同属性决定。定义了两种计算相似度的模型：差异模型和比率模型。但这些模型没有考虑属性在相似度计算时的差异性。

#### 2.2.4 基于语义关系的方法

基于语义关系的方法也被称为基于语义距离的方法，根据概念在本体层次结构中的位置来计算语义相似度[10]。Rada 认为，对于只有 is_a 关系的层次结构，任何两个节点之间有且只有一条最短路径，语义相似和语义距离是等价的[11]。但该方法只考虑了概念之间的 is_a 关系，而忽略了 part_of 等关系，计算的相似度比较粗糙。而且语义距离除了受节点间的路径长度影响外，还受其他一些因素的影响，如概念层次结构的深度等。基于语义关系的方法也可以看作是一种演绎的方法[12]。

## 3 基于语义距离的概念相似度度量

### 3.1 定义语义距离

定义 1 中的 $H$ 关系只描述了一部分语义关系，除此以外，我们更关心的是概念之间的语义关联。概念间的语义关联可以由语义距离来反映。下面具体讨论语义距离的计算方法。

基于本体的语义距离主要测量本体中概念间连接边的长度[13]，概念间的语义关联程度通过几何度量来表征。两个概念在本体中的连接路径越短，它们就越相似，每条连接边的长度由其包含的信息量决定[14]。

设 $p(c)$ 为概念 $c$ 在整个概念集中的发生概率，设 $count(c)$ 为概念 $c$ 在本体 $O$ 中的出现次数，$count(O)$ 为本体 $O$ 中的概念总数，由于概念可能以不同抽象层次的形式出现，计算概念的总出现次数时应累加其所有子概念的出现次数。$p(c)$ 的计算方法为：

$$p(c)=\frac{count(c)+\sum_{H(c',c)} count(c')}{count(O)} \tag{1}$$

由式(1)可以看出，$p(c)$ 随着 $c$ 所在层次的上升单调增加，且 $p(Root)=1$。如图 1 所示，$p(c_0)=1$，$p(c_1)=0.5$。

设 $parent(c)$ 为概念 $c$ 在本体 $O$ 中的父亲集合，即

$$parent(c)=\{c'\mid H(c,c'),\neg\exists c'', s.t.\ H(c,c''),H(c'',c')\} \tag{2}$$

根据 $H$ 关系的无环性，易知，

$$\forall c\in C, \text{if } parent(c)\neq\phi, \text{then}\mid parent(c)\mid=1 \tag{3}$$

其中，$\mid parent(c)\mid$ 表示集合 $parent(c)$ 中包含的元素个数。

由式(3)可知，如果 $c$ 存在父亲，则其父亲是唯一的。但是，在一些本体图中，一个概念可能存在多个父亲，在这种情况下，可以先将本体图转化为本体层次树[15]，然后再进行相应的计算。关于这个问题，在此不再进行过多的讨论。

根据信息论可知，如果一个概念出现的频率越大，它所包含的信息量就越少，反之越小，则越多。因此，概念 $c$ 所包含的信息量为：

$$I(c)=-\log(p(c)) \tag{4}$$

连接边 $c\rightarrow parent(c)$ 包含的信息量为

$$I(c\rightarrow parent(c))=-\log(p(c\rightarrow parent(c)))=$$
$$-\log(p(c\mid parent(c)))=-\log\left(\frac{p(c)}{p(parent(c))}\right)=$$
$$I(c)-I(parent(c)) \tag{5}$$

因此，连接边 $c\rightarrow parent(c)$ 的长度正比于它所包含的信息量：

$$length(c\rightarrow parent(c))\propto I(c\rightarrow parent(c)) \tag{6}$$

由式(5)和式(6)可得：

$$length(c\rightarrow parent(c))=\beta\cdot(I(c\rightarrow parent(c))) \tag{7}$$

其中，$\beta$ 为一个常量，是一个比例因子。为了表达方便，在本文中，令 $\beta=1$，得到：

$$length(c\rightarrow parent(c))=I(c\rightarrow parent(c)) \tag{8}$$

由于 $\forall c\in C, H(c,Root)$ 成立，给定任意两个概念 $c_1$ 和 $c_2$，设 $lca(c_1,c_2)$ 为 $c_1$ 和 $c_2$ 在本体 $O$ 中的最小公共祖先(Least Common Ancestor)，最小共同祖先 $lca(c_1,c_2)$ 是 $c_1$ 和

$c_2$ 共同祖先路径上最大深度的概念节点。对于任意两个概念节点，一定存在一个最小公共祖先，如图 1 所示，$lca(c_3,c_9)=c_1$，$lca(c_1,c_5)=c_0$，即本体的根概念节点(*Root*)，因此，

$$lca(c_1,c_2)=\{c\mid H(c_1,c),H(c_2,c),\neg\exists c',s.t.\ H(c_1,c'),H(c_2,c'),H(c',c)\} \tag{9}$$

若 $c_1$ 和 $c_2$ 满足 $H(c_1,c_2)$，则 $c_1$ 到 $c_2$ 的父子链为 $pcl(c_1,c_2)$，则

$$pcl(c_1,c_2)=\{c_1^0,c_1^1,c_1^2,\cdots,c_1^{i-1},c_1^k\mid c_1=c_1^0,c_2=c_1^i,i\geqslant 1,\forall k,0\leqslant k\leqslant i-1,c_1^{k+1}=parent(c_1^k)\} \tag{10}$$

**定义 2**(连接路径) 根据 $\forall c\in C,H(c,Root)$ 以及 $H$ 关系的无环性，$c_1$ 和 $c_2$ 在本体 $O$ 中的连接路径有且只有一条，记作 $path(c_1,c_2)$，而 $c_1$ 到 $lca(c_1,c_2)$ 的路径与 $c_2$ 到 $lca(c_1,c_2)$ 的路径的并集即为这样的连接路径，定义为：

$$path(c_1,c_2)=pcl(c_1,lca(c_1,c_2))\cup pcl(c_2,lca(c_1,c_2)) \tag{11}$$

**定义 3**(语义距离) 语义距离反映了两个概念的语义相似程度。基于定义 2 及上述公式，将语义距离定义为：

$$\begin{aligned}sem\_dis(c_1,c_2)&=\sum_{c\in\{path(c_1,c_2)-lca(c_1,c_2)\}}length(c\rightarrow parent(c))\\&=\sum_{c\in\{path(c_1,c_2)-lca(c_1,c_2)\}}I(c)-I(parent(c))\\&=(I(c_1)-I(lca(c_1,c_2))+(I(c_2)-I(lca(c_1,c_2))\\&=2\log(p(lca(c_1,c_2)))-(\log(p(c_1))+\log(p(c_2)))\end{aligned} \tag{12}$$

从式(12)中可以看出，语义距离的定义实际上包含两方面的信息，一方面，领域本体的构成决定了 $lca(c_1,c_2)$ 的位置，这在相似度度量时，把概念之间的继承关系以及概念在本体层次中所处的深度对相似度的影响均包含在了语义距离之内；另一方面，$p(lca(c_1,c_2))$，$p(c_1)$，$p(c_2)$ 的值来自于概念集的统计信息。相对于不同的领域本体，概念间的语义距离是不同的，而不同的概念相对于同一个领域本体，其语义距离也是不同的。

这种计算方法综合了概念间的语义关系以及客观发生的统计信息，有助于更准确地模拟客观世界的原貌。

### 3.2 计算语义相似度

得到两个概念之间的语义距离之后，需要构造合理的语义距离到语义相似度的转换函数，从而将语义距离转化为语义相似度。语义相似度函数应满足如下几个主要特性[16]：(1)语义相似度函数的输出必须在[0,1]区间内；(2)当语义距离为 0 时，语义相似度为 1。即当两个概念相同时，语义相似度为 1；当语义距离为无穷大时，语义相似度为 0；(3)语义相似度随语义距离的增加而减小。即语义距离大的概念间的相似度小于语义距离小的概念间的相似度。

从理论上分析，满足以上 3 个主要特性的语义相似度转换函数有很多，简单、易用的可以分为以下 3 种：

$$sem\_sim1=1/(\rho*sem\_dis+1) \tag{13}$$

$$sem\_sim2=1/(\rho*sem\_dis^2+1) \tag{14}$$

$$sem\_sim3=1/\rho*e^{sem\_dis} \tag{15}$$

在以上 3 个公式中，$\rho$ 表示调节因子，在具体计算概念之间的语义相似度时，可以通过调整 $\rho$ 值来确定不同系统需要的相似度。一般情况下，取 $\rho=1$。以上 3 个转换函数反映了语义距离和语义相似度之间呈不同递减速度的反比关系。由于式(14)和式(15)中函数的下降速度过快，当本体层次数较大时，不能较好地反映出本体概念之间的语义相似度，因此，我们选取式(13)作为语义距离到语义相似度的转换函数。

## 4 实验评价

为了考察本文所提方法的有效性，我们与其他几个经典的相似度计算方法进行了对比分析。在该实验中，采用的转换函数为式(13)，其中，取 $\rho=1$。实验对象是基于 WordNet 的一个简单本体的一部分，如图 2 所示。

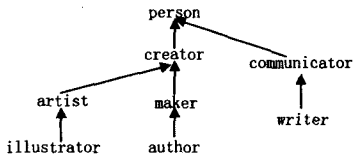


图 2 基于 WordNet 的简单本体

采用不同的相似度度量方法进行度量，可以得到各概念之间不同的相似度，如图 3 所示。

(a) Synonymy similarity

| | illustrator | author | creator | person | writer |
|---|---|---|---|---|---|
| illustrator | 1.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| author | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| creator | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| person | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |
| writer | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |

(b) Gloss overlap

| | illustrator | author | creator | person | writer |
|---|---|---|---|---|---|
| illustrator | 1.0 | 0.05 | 0.07 | 0.0 | 0.02 |
| author | 0.05 | 1.0 | 0.0 | 0.0 | 0.19 |
| creator | 0.07 | 0.0 | 1.0 | 0.06 | 0.02 |
| person | 0.0 | 0.0 | 0.06 | 1.0 | 0.04 |
| writer | 0.02 | 0.19 | 0.02 | 0.04 | 1.0 |

(c) Upward cotopic similarity

| | illustrator | author | creator | person | writer |
|---|---|---|---|---|---|
| illustrator | 1.0 | 0.37 | 0.43 | 0.4 | 0.18 |
| author | 0.37 | 1.0 | 0.43 | 0.29 | 0.36 |
| creator | 0.43 | 0.43 | 1.0 | 0.4 | 0.18 |
| person | 0.4 | 0.29 | 0.4 | 1.0 | 0.25 |
| writer | 0.18 | 0.36 | 0.18 | 0.25 | 1.0 |

(d) The proposed method

| | illustrator | author | creator | person | writer |
|---|---|---|---|---|---|
| illustrator | 1.0 | 0.42 | 0.47 | 0.53 | 0.36 |
| author | 0.42 | 1.0 | 0.47 | 0.53 | 0.36 |
| creator | 0.47 | 0.47 | 1.0 | 0.83 | 0.47 |
| person | 0.53 | 0.53 | 0.83 | 1.0 | 0.53 |
| writer | 0.36 | 0.36 | 0.47 | 0.53 | 1.0 |

图 3 各种相似度度量方法的比较

如图 3 所示，(a)是 synonymy similarity[16] 的计算结果，(b) 是 gloss overlap[17] 的计算结果，(c)是 Upward cotopic similarity[18] 的计算结果，(d)是本文所提方法的计算结果。

由于不存在精确的尺度来衡量两个概念之间绝对的相似度，在概念相似度度量上，一方面可以通过一些相似度计算方法得到具有一定参考价值的相似度值，对比不同方法的性能优劣；另一方面可以通过领域专家根据经验和常识来判断计算结果的正确性。因此可以分别从上述两方面考察相似度算法的合理性及有效性。

从图 3 中的实验结果来看，synonymy similarity 方法只能发现同一概念间的相似度，这在应用中并没有太大的实际意义；gloss overlap 得到了优于 synonymy similarity 的结果，但也忽略了许多概念之间的语义相似度；Upward cotopic similarity 和本文所提方法均得到了相对理想的结果，但是 Upward cotopic similarity 得到的有些概念(如 person 和 writer)的相似值较低，不能较好地反映特定应用领域中各实体之间的语义关系。另外，我们也请语言学及计算机科学的领域专家对实验结果进行了评价，他们对本文方法的计算结果也给予了较好的评价。从图 3 所示的计算结果以及在 WordNet 中对实验所用本体概念的定义来看，本文所提方法的结

果比较接近人类的直观认识，更加符合特定领域的语义匹配。

**结束语** 本文提出了一种用于服务匹配中的本体概念之间语义相似度的计算方法，在概念相似度度量上，不但考虑了概念之间的继承关系和概念在本体中所处深度对相似度的影响，也考虑概念客观发生的统计信息。实验表明，本文的方法比较符合人类的直观认识，得到了概念间有价值的语义相似度。

在语义服务匹配中，本方法只是考虑了影响语义相似度的几个比较重要的因素，随着语义 Web 和本体驱动的信息系统研究和应用的推广，会有更多的、更复杂的语义结构信息可以利用，如本体中的实例及属性关系对概念语义相似度的影响等，如何构建考虑更多因素的语义相似度度量方法是我们下一步的工作重点。

## 参 考 文 献


[1] Giunchiglia F, Yatskevich M, Shvaiko P. Semantic Matching: Algorithms and Implementation[M]. Journal on Data Semantics IX, Springer Berlin / Heidelberg, 2007: 1-38

[2] Burstein M, Hobbs J, Lassila O, et al. OWL-S: Semantic Markup for Web Services[OL]. http://www.daml.org/services/owl-s/1.1/overview, 2004

[3] Budanitsky A, Hirst G. Evaluating wordnet-based measures of lexical semantic relatedness [J]. Computational Linguistics, 2006, 32(1): 13-47

[4] Pease A. Standard Upper Ontology Knowledge Interchange Format[OL]. http://suo.ieee.org/, 2000

[5] Gruber TR. A Translation Approach to Portable Ontology Specifications[J]. Knowledge Acquisition, 1993, 5(2): 199-220

[6] Cohen W, Ravikumar P, Fienberg S. A comparison of string metrics for matching names and records [C] // Proceeding of KDD Workshop on Data Cleaning and Object Consolidation. 2003: 73-78

[7] Maynard DG, Ananiadou S. Term extraction using a similarity-based approach [A] // Bourigault D, Jacquemin C, L'Homme MC, eds. Recent Advances in Computational Terminology[C]. John Benjamins, 1999: 261-278

[8] Cerbah F, Euzenat J. Traceability between models and texts through terminology [J]. Data and Knowledge Engineering, 2001, 38(1): 31-43

[9] Tversky A. Feature of similarity [J]. Psychological Review, 1977, 84 (4): 327-352

[10] 吴健，吴朝晖，李莹，等. 基于本体论和词汇语义相似度的 Web 服务发现[J]. 计算机学报，2005，28(4)：2054-2062

[11] Rada R, Mili H, Bicknell E, et al. Development and application of a metric on semantic nets [J]. IEEE Transaction on System, Man, and Cybernetics, 1989, 19(1): 17-30

[12] Giunchiglia F, Shvaiko P, Yatskevich M. Discovering missing background knowledge in ontology matching[C] // Proceedings of 16th European Conference on Artificial Intelligence (ECAI). 2006: 382-386

[13] Jiang J, Conrath D. Semantic similarity based on corpus statistics and lexical taxonomy[C] // Proceedings of the International Conference on Research in Computational Linguistics, 1997: 19-33

[14] 杨立，左春，王裕国. 基于语义距离的 K-最近邻分类方法[J]. 软件学报，2005，16(12)：2054-2062

[15] 梁敏，郭新涛，阮备军，等. X_Dist——一个柔性语义距离函数[J]. 计算机研究与发展，2004，41(10)：1728-1736

[16] Giunchiglia F, Shvaiko P, Yatskevich M. S-Match: an algorithm and an implementation of semantic matching[C] // Proceedings of 1st European Semantic Web Symposium (ESWS). 2004: 61-75

[17] Lesk M. Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone [C] // Proceedings of 5th Annual International Conference on Systems Documentation (SIGDOC). 1986: 24-26

[18] Madche A, Zacharias V. Clustering ontology-based metadata in the semantic web[C] // Proceedings of 6th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD). 2002: 348-360


(上接第 180 页)

从测试结果可以看出，对于激光与等离子体相互作用模拟的电场强度物理量使用 gzip 无损压缩的压缩比仅为 1.07，小的压缩比使得无损压缩变得没有意义。使用内插预测算子 $R$ 的有损压缩可以取得比较理想的压缩比，而且比较使用外插预测算子 $L^2$ 的压缩比提高约 17%，同时压缩时间开销减少，解压缩时间开销增加，总的时间开销没有明显增加。

**结束语** 我们研究比较了内插预测与外插预测算子，改进了内插预测算子的反预测时间开销比较大的问题，使用外插预测实现了纯量场数据的无损压缩，使用内插预测实现了纯量场数据的有损压缩。本文提出的压缩方法的突出优点是内存开销比较小，适合于大规模数值模拟产生的纯量场数据的压缩。使用光滑的数学模拟测试数据和真实的物理模拟数据进行的测试实验表明，本文方法取得了比较好的效果，具有较高的应用价值。

## 参 考 文 献


[1] Fowler J, Yagel R. Lossless compression of volume data[C] // Symposium on Volume Visualization. 1994: 43-50

[2] Lawrence I, Peter L, Jarek R. Out-of-core compression and decompression of large n-dimensional scalar fields [J]. EUROGRAPHICS, 2003, 22(2): 89-97

[3] Vadim E, Dag F, Peter F. Lossless Compression of high volume data from simulation[A] // IEEE Computer Society[C]. 2000: 754-765

[4] Chiueh T, Yang C, He T, et al. Integrate volume compression and visualization[C] // Visualization'97. 1997: 329-336

[5] Aaron T, Robert M, John M. Wavelets applied to lossless compression and progressive transmission of float point data in 3D curvilinear data[M]. IEEE Computer Society Press, 1996: 385-388

[6] Manuel N. Gamitoa and Miguel Salles Dias, Lossless Coding of Floating Point data with JPEG 2000 Part 10, 2004

[7] 吴国清，陈虹. 基于最优内插预测的科学数据压缩方法[J]. 计算机科学，2007，34(8)：15-17

[8] 吴国清，陈虹. 一种科学数据无损压缩方法[J]. 计算机工程与应用，2006，42(5)：172-175

[9] Lawrence I, Peter L. Predictors for streaming compression of scalar fields[J]. EUROGRAPHICS, 2006, 25(3): 1077-1084

[10] Peter L, Martin I. Fast and efficient compression of floating point data[J]. IEEE Transactions on Visualization and Computer Graphics, 2006, 12(5): 869-875

[11] 李庆扬，关治，白峰杉. 数值计算原理[M]. 北京：清华大学出版社，2002：185-191