

基于模板的服务选择方法

周云, 贺毅辉, 刘志忠, 杨楠

(中国人民解放军理工大学 指挥自动化学院, 江苏 南京 210007)

摘要:在 Web 环境中, 服务是千变万化的, 满足相同功能的服务可能会有多个, 这些服务具有不同的 QoS 参数, 如何从中选择满足模板中各服务节点功能需求的具体服务, 形成一个可执行的组合服务流程来完成用户的需求, 成为服务组合中的一个关键问题。在基于模板的服务组合基础之上, 提出一种解决模板选择和服务选择过程中的 QoS 最优化问题的模型, 建立模板 QoS 模型, 保证最大概率选择到 QoS 最优模板。在模板的节点选择服务过程中基于改进多目标遗传算法, 同时优化多个 QoS 参数, 最终产生一组满足约束条件的 QoS 最优服务组合流程集。实验结果表明了算法具有良好的可行性和有效性。

关键词:服务组合; 模板; 双重 QoS 保障机制; QoS 全局最优; 多目标遗传算法

中图分类号:TP31

文献标识码:A

文章编号:1673-629X(2012)06-0114-05

Method of Web Service Selection Based on Template

ZHOU Yun, HE Yi-hui, LIU Zhi-zhong, YANG Nan

(Institute of Command Automation, PLA University of Science and Technology, Nanjing 210007, China)

Abstract: Services are protean in Web environment. There may be several services which provide the same function, and those services have different quality of service. Selecting a service instance for each service node of the template with optimal QoS under the condition of satisfying the global QoS constraints, is the key of service composition. Based on composition template, proposed a model to solve the problem of QoS global optimal in the process of template selection and service selection. Build the model of template QoS to maximize the probability to select optimal QoS template. The problem of service selection with QoS global optimal is transformed into a multi-objective services composition with QoS constraints. Produce a set of optimal services composition process with constraint principle by means of optimizing various QoS parameters simultaneously. Experimental results show the feasibility and efficiency of this method.

Key words: service composition; template; double QoS protection mechanism; QoS global optimal; multi-object genetic algorithm

0 引言

服务组合源自于软件重用, 根据客户的实际需求, 以给定的应用逻辑将已有的服务组织到一起, 从而创建出新的或更高质量的应用。而服务选择是指依次为组合服务流程中的每个抽象服务, 从其满足条件的实体服务集中选择一个候选服务, 组成一个组合服务实例。

服务组合主要根据对 Web 服务研究的侧重点来划分类别: 根据组合的自动化程度将服务组合分为手动、半自动和自动三大类; 根据对参与组合的 Web 服务的绑定时间, 将其分为静态的和动态的两大类; 根据所依赖技术分为过程驱动的服务组合和语义驱动的服务组合^[1]。

基于模板的服务组合方法, 是一种半自动的动态服务组合方法, 该方法允许服务组合设计者为某些业务活动设定服务模板, 而不需指定到具体的服务^[2]。在服务模板中对希望的目标服务进行描述, 内容包括目标服务的输入、输出、服务质量等。在执行之前, 将实体服务绑定到事先制定好的服务模板中, 从而实现服务的组合。在过去的研究中, 基于模板的组合方法, 很少考虑到组合过程中的 QoS 优化问题, 缺乏相对完善的 QoS 机制来保障模板用户的服务质量需求。

文中提出了一种基于模板的 QoS 最优服务组合方法, 该方法采用双层 QoS 保障机制: 在 QoS 服务层, 把模板的服务选择问题转化为一个带多约束条件的多目标服务组合问题, 通过利用多目标遗传算法的智能优化原理^[3], 同时优化各个 QoS 目标参数, 产生一组满足约束条件、QoS 最优的服务组合流程集, 用户可以根据自身需求从中选择一组满意的组合流程; 在 QoS 模板层, 通过记录模板的执行效果以及用户的评价, 形

收稿日期: 2011-11-26; 修回日期: 2012-02-27

基金项目: 江苏省自然科学基金项目 (BK2010130)

作者简介: 周云 (1985-), 男, 硕士研究生, 研究方向为系统集成与优化; 贺毅辉, 教授, 研究方向为作战仿真、决策支持。

成模板的 QoS 目标参数,然后对满足功能需求的模板通过求解 QoS 最优模板的方式进行排序,最后根据用户的需求,从排名靠前的模板中选择合适个体进行下一步服务选择工作。

1 模板的表示及相关定义

服务组合模板 (Service Composition Template, SCT) 是由抽象的服务和一定的逻辑单元组成的复合结构^[4]。模板主要是根据领域专家的知识 and 经验,通过对领域需求的分析,结合行业背景、业务规则,归纳总结出一些规则性的工作模式。利用基本的服务和流程结构单元,构造出可重用的业务过程框架。用户可以根据实际需求对这样的流程框架进行直接套用、裁剪和扩充,从而构建应用^[5]。

定义 1:服务组合模板是一个五元组:

$SCT = \langle TemId, TemDes, SSet, TemTs, TemQoS \rangle$

TemId 是模板的唯一标识,用户用以识别不同的服务模板信息,它包括多条信息项:模板的名称、模板的提供者、模板的版本信息和模板发布时间;TemDes 是对模板的描述,描述模板涉及的领域、输入/输出信息、服务功能以及模板类型;ServiceSet 描述组成此模板的服务集,它包括服务集合 Services 以及服务关系集合 Relations;TemTs 表示模板引用次数;TemQos 是对模板的整体 QoS 描述。

定义 2:服务节点 (Service Node, SN)。将模板中的抽象服务定义为服务节点,它是构成服务模板的基本逻辑单元,只包括功能和接口信息的描述,不指向具体的服务。

定义 3:服务类群 (Service Class, SC)。服务类群是指具有相同功能和相同的调用接口的由不同的服务提供者发布的一组服务集。服务类群中各个服务的 Qos 属性各不相同。有的服务可能同时隶属于多个服务类群。

定义 4:模板实例 (SCT Instance, SCTI)。将模板中的每一个抽象服务实例化,进而形成满足功能需求

的由实体 Web 服务构成的组合服务。

图 1 是一个服务组合模板的基本模型,模板由服务节点和业务逻辑关系组成。在模板执行时,通过搜索出满足各服务节点功能描述相匹配的服务集群,并结合实际非功能性需求从中选择绑定实体服务,形成一个模板实例执行相应的任务,返回用户所需结果。

2 服务选择模型

2.1 服务 QoS 五维模型

文中对服务引入 QoS 保障机制,主要参考文献 [5] 中的 AgFlow 提出的五维 QoS 模型,即考虑服务的执行费用、响应时间、服务声望、可靠性和可用性。

其中服务的执行费用 (Price, P) 为调用该服务所应支付的费用;服务的响应时间 (Time, T) 定义为从一个用户发起服务请求到收到服务响应之间的时间间隔;服务的声望 (Reputation, Rep) 定义为一个服务提供者受到好评的比率;服务的可靠性 (Reliability, R) 定义为其被成功调用的概率;服务的可用性 (Availability, A) 定义为其正常运行的概率。

2.2 服务组合中的基本组合模型及 QoS 计算

2.2.1 基本结构及其 QoS 计算

构成模板的服务组合流程包含四种基本的结构:顺序结构 (Sequence)、选择结构 (Choice)、并行结构 (Parallel) 和循环结构 (Loop),如图 2 所示。大部分服务组合模板内的服务组合流程模型,都可以用这四种基本结构组合而成。

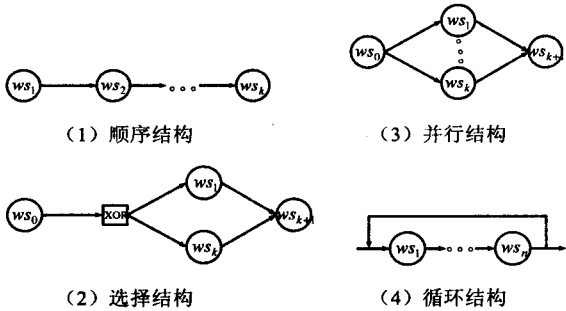


图 2 服务组合流程基本结构

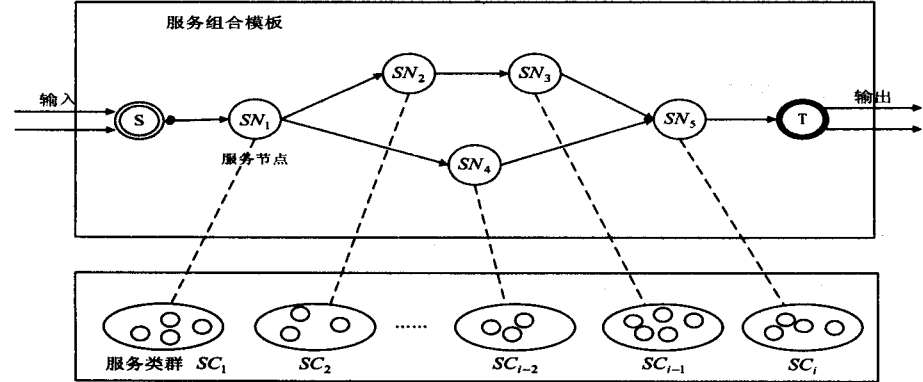


图 1 服务组合模板模型

文中提出的 5 种 QoS 参数,在各个基本组合结构中的计算方法有所不同。文献 [6] 给出了具体的计算函数,但是并没有给出全局路径的合成方法,文中对于合成方法不进行讨论。

2.2.2 目标函数和约束函数的确定

基于模板的 QoS 最优服务选择就是通过模板的

实例化,使得到的模板实例满足一定的可靠性和可用性,响应时间和执行费用达到最优。由此可以确定目标函数和约束函数^[7,8]:

目标函数:

$$\min F(X) = (T(X), P(X)) \quad (1)$$

约束函数:

$$R(X) \geq R_0; A(X) \geq A_0; \text{Rep}(X) \geq \text{Rep}_0 \quad (2)$$

其中 $T(X)$ 、 $P(X)$ 、 $\text{Rep}(X)$ 、 $R(X)$ 、 $A(X)$ 分代表模板中 QoS 参数响应时间、执行费用、声望、可靠性、可用性, R_0 、 A_0 、 Rep_0 分别代表服务组合过程中所要求的最小可靠性、最小可用性和最小声望。式 1、式 2 表示要在可靠性、可用性和声望满足约束的条件下,其余参数同时极小化,来满足模板整体 QoS 最优化。

2.3 基于多目标遗传算法的模板服务选择算法

模板服务选择问题属于 NP-Complete 问题^[9],不能应用传统的优化求解技术求解,文中提出一种基于多目标遗传算法的 QoS 全局最优组合模板服务选择算法,把 QoS 全局最优的模板服务选择问题转化为一个多目标服务组合优化问题。

2.3.1 算法过程

1) 算法参数设定:种群规模 N , 进化代数 T , 交叉概率 p_1 , 变异概率 p_2 ;

2) 初始化: $t = 0$, 初始化种群 P_1 , 外部种群 P_2 ;

3) 进行个体适应值的分配和多样性的保持;找出最优解,判断算法是否达到收敛条件,若满足,终止算法,否则继续执行;

4) 利用种群 P_1 和外部种群 P_2 进行优良解的保持;

5) 个体选择:根据个体的排序值,采用轮盘赌方式从种群 P_1 中选择优势个体,与外部种群混合后赋给交配池;

6) 按照策略执行交叉和变异操作,产生新的染色体作为新的种群;

7) 返回到步骤 3) 直到获得最优解集。

2.3.2 编码

算法的编码方式是将模板的每一个具体的模板实例编码为一个染色体。对模板的各个服务节点进行排序,采用整数定长编码的方式。每一个整数组长度为模板包含的服务节点数,数组中的元素的顺序为服务节点的顺序,每个元素的值为对应所选择的具体的服务的编号。

2.3.3 初始化操作

种群初始化是遗传算法的重要步骤,其结果直接影响算法的收敛速度。具体过程:设定种群规模 N , 输入约束函数集 $F = \{f_i \mid 1 < i < m\}$, 输出为初始种群 P_1 。从所有染色体中,随机选择一条路径,如果满足约

束条件,则加入到 P_1 中,否则继续,直到满足 $|P_1| = N$ 为止。

2.3.4 个体适应值分配和多样性保持

适应值分配,就是给每一个染色体赋予一个标量适应值,对当代种群内的个体进行排序。排序规则如下:针对不同的目标,分别对各个个体进行排序,即针对第 i 个目标函数对种群 P_1 中的 N 个个体按照其目标函数值的优劣进行排序,得到排序序列 \vec{X}_i , 由此得到 m 个这样的序列。那么根据个体的排序计算其适应度:

$$\text{Fit}_i(X_j) = \begin{cases} (N - R_i(X_j))^2 & R_i(X_j) > 1 \\ KN^2 & R_i(X_j) = 1 \end{cases} \quad i = 1, 2, \dots, m$$

$$\text{Fit}(X_j) = \sum_{i=1}^m \text{Fit}_i(X_j), \quad j = 1, 2, \dots, m \quad (3)$$

其中 N 为个体总数; m 为目标函数总数; X_j 为种群中的第 j 个个体; R_i 为个体在种群中对第 i 个目标函数值优劣排序的序号; $\text{Fit}_i(X_j)$ 表示 X_j 对目标 i 所得的初始适应值; $\text{Fit}(X_j)$ 表示个体 X_j 的综合初始适应值; k 为 $(1, 2)$ 区间的一个实数,用于加强最优个体的适应值。由上式可以看出,对总体表现较优的个体往往得到较好的适应值,获得更多参与竞争的机会。

在很多问题中遗传算法往往会因为丧失群体的多样性而陷入“早期收敛”的困境,为了克服这种问题,保持群体的多样性,目前大多数研究采用的是小生境技术,文中主要采用香农的信息熵理论来保持群体的多样性。群体 P_1 中个体 X_i , 对任意给定的个体 $X_j \in P_1$, 所包含的节点相同率为 p_{ij} , 则信息熵:

$$H_{ij} = -p_{ij} \ln(p_{ij}) \quad (4)$$

由此可以计算出 X_i 和 X_j 的相似度 $S_{ij} = 1/(1 + H_{ij})$ 。给定 ω ($0.9 \leq \omega \leq 1$), $n = 0$, 如果 $S_{ij} > \omega$, 则 $n = n + 1$ 。 X_i 在群体中的密度 $D_i = n/N$ 。则 X_i 的适应度:

$$\text{fit}(X_i) = \text{Fit}(X_i)/D_i \quad (5)$$

从公式 5 中可以看出,群体中初始适应值较小、密度较大的个体,将获得较小的适应度,这样就防止进化过程中,相似的个体充斥整个群体,保持了群体的多样性,从而保证了良好的进化方向。

2.3.5 优良解保持

由于遗传算法具有一定的随机性,往往导致进化过程中产生的优良解丢失,因此,为了保证算法的优化性和收敛性,引入了外部群体 P_2 对当前产生的优良解实施保留。在进化过程的每一代,保存满足约束条件的最优解到外部群体。主要实施方法:把第一代进化产生的满足约束条件最优解加入到 P_2 中采用公式 3 的方法进行排序,把排名最后的个体从 P_2 中移除。以后每一代进化得到的最优解与最新的 P_2 进行比较替换,

这样就保证了算法结束时,得到了最好的解。理论证明,引入这种优良解保持的算法将收敛到最优解^[10]。

2.3.6 交叉与变异

文中采用两点交叉策略,相邻的两个染色体,随机选择一个交叉节点依概率 p_1 进行交叉。

变异操作是在染色体上自发地产生随机的变化,其中,变异位的操作是在某个元素对应的服务类群中随机选择新的 Web 服务替换当前的服务。通过交叉和变异,产生新的个体加入到种群中。

3 模板 QoS 模型

服务组合模板很好地封装了领域内业务服务及其组合逻辑,使得用户使用时无须自己再重新组合这些服务。随着服务模板应用规模的不断扩大,越来越多的模板得以构建,与 Web 服务一样,网络上可能存在着相当一部分功能和接口相同相似的模板,由于出自不同开发者,其内部结构的不同随之带来的是效率、可靠性等方面的好坏优劣。如果在第一时间能够判断出 QoS 最优的服务组合模板,整体服务质量也将达到最优。基于这种考虑,引入了模板的 QoS。

3.1 模板 QoS 参数

服务组合模板是由若干的服务节点按照一定的序列组合而成,这些服务节点是抽象的服务,只有执行时才绑定到具体的服务。所以模板只有实例化以后才能计算出其 QoS 属性,然而该属性只能代表当前模板实例的 QoS,并不能代表模板的 QoS。尽管如此,一个模板所提供的服务质量可能与具体的服务有关,但是模板不同的内部结构对其提供的服务是有一定影响的。比如说功能相同的两个不同模板,其内部流程的合理性影响了它的执行效率;其次,虽然服务是动态变化的,但在一定时间内,一个模板所生成的模板实例的 QoS 是相对平稳的。

文中提出的模板 QoS 模型,与服务的 QoS 类似,主要考虑到模板的执行时间、执行费用、可靠性、可用性、声望,即 $TemQoS = \langle TemT, TemP, TemRep, TemR, TemA \rangle$ 。其中 $TemT$ 是指模板最近 n 次被调用的平均执行时间; $TemP$ 指模板最近 n 次被调用的平均执行价格; $TemRep$ 记录用户对模板的评价; $TemR$ 是指模板成功执行的概率。考虑到服务的不稳定性,模板的执行时间和费用在每次具体执行过程中是不相同的,文中根据模板在最近使用过程中的表现来统计平均的执行时间和执行价格。

3.2 模板选择

在一定的可靠性范围内,对模板进行选择,一种便捷的方法是对各个模板 QoS 参数进行线性加权后排序,选择最优^[11],这种方法把多目标问题转化成为单

目标问题,不能满足用户的要求。

定义 5: QoS 最优模板:给定一个多目标模板 QoS 最优问题 $\min F(T)$,称 $T^* \in \Omega$ 为最优 QoS 模板,若 $\forall T \in \Omega$,均有 $F_i(T) \geq F_i(T^*)$, ($i = 1, \dots, m$),且至少存在一个目标 i 满足 $F_i(T) > F_i(T^*)$ 。其中 \geq 和 $>$ 分别表示不劣于和优于关系。

利用定义 5 中最优 QoS 模板的定义,将符合要求的模板集中的 QoS 最优模板标记为 1,然后从竞争中去除该模板,在余下的模板集中继续寻找 QoS 最优模板标记为 2,直到所有模板都分配到次序。而后用户可以根据自身需求和爱好,选择排名靠前的模板进行下一步的服务选择工作。

4 实验及分析

通过仿真实验验证文中所提出模型获取 QoS 最优服务组合的能力及算法运行的效率,并且与当前所采用的算法进行对比。实验环境如下:100M 局域网,实验主机为 Pentium IV 2.93GHz 处理器,1G 内存,操作系统为 Windows XP, jdk1.5.0。

4.1 实验设计

服务选择实验采用图 1 中的模板模型,模板中的每一个节点对应一个服务类群。服务群的建立和维护文中不作描述。实验采用文献[12]中的合成执行路径全局 QoS 计算模型,对模板的全局 QoS 进行计算。

4.2 实验参数选择

多目标遗传算法的参数主要包括种群大小 P 、交叉率和变异率。文献[4]证明了:满足以下三个条件的遗传算法能收敛到最优解:

- (1) 交叉概率 $p_1 = 1$;
- (2) 变异概率 $p_2 = 0.15$;
- (3) 每一代群体在选择操作前进行最优解的保持。

文中采用该策略实现服务选择。

各个服务的 QoS 参数采用随机生成的方式在一定范围内生成。具体设置为执行时间 T 为(0,10),单位为秒;费用 C 为(0,50),单位为元;声望 Rep 为(0,10);可靠性 R 为(0,1);可用性为(0,1)。其中,约束条件最小声望为 4,最小可靠性为 0.2,最小可用性为 0.2。

4.3 实验结果分析

4.3.1 有效性实验

实验考虑种群规模 5、10、15,进化代数 T 取 100、200、300 和 400 的情况下算法的开销。表 1 中给出了各种不同取值下算法的时间开销。

从实验结果可以看出,影响算法效率的两个主要因素是进化代数和种群的规模。在相同的进化代数

下,随着种群规模的增加,算法运行时间并没有大量增加。并且在种群规模为 15 的情况下,运行时间为 11 秒,这个结果是可以接受的。

表 1 算法平均执行时间(s)

<i>P</i>	<i>T</i> = 100	<i>T</i> = 200	<i>T</i> = 300	<i>T</i> = 400
<i>P</i> = 5	1.8	2.8	4.0	5.8
<i>P</i> = 10	2.2	3.6	5.3	9.7
<i>P</i> = 15	2.4	5.2	6.8	11.3

4.3.2 可行性实验

可行性实验的目的,是验证算法寻找到 QoS 最优组合方案的可行性。实验首先用穷举法计算出最优解集,然后用文中的算法在相同条件下计算出的最优解集与其进行比较,计算出文中算法获取到各目标 QoS 属性最优解的概率。表 2 中的数据是通过算法运行 10 次所得结果。其中,取 *T* = 400。

表 2 最优解比例(%)

<i>P</i>	Time	Price
<i>P</i> = 5	100%	100%
<i>P</i> = 10	95%	90%
<i>P</i> = 15	100%	95%

从实验结果可以看出,在取进化代数为 400 的情况下,找到各项指标的最优值的概率都在 90% 以上,这个结果是比较理想的,算法具有可行性。

5 结束语

Web 服务以其独特的优势逐步获得大多数人的青睐,服务组合 QoS 的不断成熟,也使得 Web 服务由一种技术转化为人们应用的工具。文中主要采用基于模板的服务组合方法,在组合过程中应用多目标遗传算法进行服务的选择,使得在保障全局 QoS 最优的情况下,算法效率得到有效的提高。实验结果证明了算法的有效性和可行性。现有服务选择算法的容错、失效恢复机制还不够完善。下一步工作中将开展服务选

择算法的自适应研究,在选择算法中融入一定的容错机制、协商机制、失效恢复机制和模板重构机制等以提高服务组合的稳定性和健壮性。

参考文献:

[1] 倪晚成,刘连臣,吴 澄. Web 服务组合方法综述[J]. 计算机工程,2008,34(2):79-81.

[2] 胡海涛,林碧英. 一种基于组合模板的大粒度服务组合方法[J]. 华北电力大学学报,2006,33(5):81-84.

[3] 郑金华. 多目标进化算法及应用[M]. 北京:科学出版社,2007:2-16.

[4] Casati F, Ilnicki S, Jin Lijie, et al. Adaptive and Dynamic Service Composition in eFlow[C]//Proc. of the International Conference on Advanced Information Systems Engineering. Stockholm:Springer-Verlag,2000:13-31.

[5] Zeng Liangzhao, Benatallah B, Ngu A H, et al. Qos-aware middle-ware for web service composition[J]. IEEE Trans. on Software Eng. ,2004,30(5):311-327.

[6] Canfora G, Dipenta M, Esposito R. An approach for QoS-aware service composition based on genetic algorithms[C]//Proc. of the 2005 Conference on Genetic and Evolutionary Computation. Washington:[s. n.],2005:1069-1075.

[7] Cui X X, Lin C. A constrained quality of service routing algorithm with multiple objectives[J]. Journal of Computer Research and Development,2004,41(8):1368-1375.

[8] 莫振华,蔡鸿明,姜丽红. 基于遗传算法的多 Qos 约束服务选择[J]. 计算机应用与软件,2009,26(3):4-6.

[9] Garey M R, Johnson D S. Computers and Intractability: A Guide to the Theory of NP-completeness[M]. New York:W. H. Freeman and Company,1979.

[10] 张建萍,刘希玉,谭业武. 改进免疫遗传算法及其应用研究[J]. 计算机技术与发展,2008,18(10):166-169.

[11] Zhao J F, Xie B, Zhang L, et al. A web services composition method supporting domain feature [J]. Chinese Journal of Computers,2005,28(4):731-738.

[12] 任开军,宋君强,肖 依,等. 基于迭代 Bargaining 策略优化服务合成执行路径[J]. 国防科技大学学报,2009,31(1):80-85.

(上接第 113 页)

2003,9(4):353-364.

[7] Avidan S, Shamir A. Seam carving for content-aware image resizing[J]. ACM Trans on Graphics,2007,27(3):10-18.

[8] 宋静娴. 基于内容的图像缩放算法研究[D]. 杭州:杭州电子科技大学,2009.

[9] 李 冉,谭光兴,宁胜花. 基于内容感知的图像缩放研究[J]. 广西工学院学报,2010,21(3):33-37.

[10] 陈剑辉. 内容敏感的图像缩放[D]. 杭州:浙江大学,2011.

[11] 王会千,杨高波,张兆扬,等. 结合显著度图和自适应能量线的快速图像缩放[J]. 计算机应用研究,2010,27(9):3594-3597.

[12] Schneiderman H, Kanade T. A Statistical Model for 3D Object Detection Applied to Faces and Cars [C]//Proc. of IEEE Conference on Computer Vision and Pattern Recognition. [s. l.];IEEE,2000.