

静态灰度图像中的人脸检测方法综述^{*}

Survey on Face Detection Methods in Gray-level Still Images

唐 伟 陈兆乾 吴建鑫 周志华

(南京大学计算机软件国家重点实验室 南京210093)

Abstract In recent twenty years, the technique of face detection and face recognition, as one of the important research area of computer vision and image understanding, attracts more and more attention. In general, face detection in gray-level still images is more difficult than that in color images. Therefore this paper briefly surveys this area and indicates some issues for exploration.

Keywords Face detection, Eigenface, Neural network, Feature extraction, Deformable template

1. 引言

在日常生活中,人的脸部特征给我们提供了大量丰富的信息。对于人脸的研究因其在身份验证、档案管理和可视化通讯等方面的巨大应用前景,备受研究者关注,成为一个非常活跃的研究领域。很多人脸研究工作都是假定图像中的人脸已经被检测和定位,而为了设计出自动人脸识别系统,快速而高效地检测人脸是需要解决的一个关键问题。

目前人脸检测还是一个相当困难的工作,其中所面临的问题可以归结如下:

①图像中是否存在人脸:这是人脸检测不同于其它有关人脸研究工作的地方,即如何判断图像中是否存在人脸,如何区别人脸和类似人脸的非人脸图像。

②检测不同表现形式的人脸:人脸可能以不同视角出现在图像中,也可能被某些物体遮挡,造成某些用于检测而需提取的人脸特征不可见。

③图像中存在着噪声:由于成像时亮度、对比度等因素的影响使图像不清晰,人脸与背景区别不大,给检测带来一定的难度。

④人脸自身的因素:由于人脸结构复杂,某些局部特征具有随机性(如眼镜、胡须、发型等),而且还存在着不同表情的人脸,这些都给人脸检测带来难度。

由此可见,人脸检测是一个复杂而具有挑战性的问题。Yang等^[1]将人脸检测定义为:任意给定一个图像或者一组图像序列,人脸检测的目的就在于判定该图像或图像序列中是否存在人脸。如果存在,则返回其位置和空间分布。根据这个定义可知,人脸检测可分为两类:第一类是在静止图像中判断是否存在人脸,若存在,则定位人脸的位置;第二类是在视频图像序列中判断是否存在人脸,若存在,则动态地跟踪人脸。显然,后一种所采用的人脸检测方法要比前一种困难。另外,根据人脸检测所处理图像的不同又可将其分为两类:基于灰度图像的人脸检测和基于彩色图像的人脸检测。对于灰度图像,目前存在的方法主要可分为三类:基于表象(appearance-based)的方法,基于特征(feature-based)的方法和基于模板(template-based)的方法。对于彩色图像,由于能够提供更丰富的信息,因此可以采用有别于灰度图像的处理方法,利用彩色图像的自身特性进行人脸检测。例如 Yang 和 Ahuja^[2]在

1998年的工作中,利用人的皮肤颜色建立了一个皮肤模型。首先根据图像的颜色进行分割,将其划分为不同的区域,然后将符合皮肤颜色的区域进行合并,直到形成椭圆区域(近似人脸轮廓)来判定人脸是否存在。

本文后续部分,将详细讨论静态灰度图像中的人脸检测方法。首先介绍基于表象的方法,然后介绍基于特征的方法和基于模板的方法,最后提出一些人脸检测领域值得关注的问题。

2. 基于表象的人脸检测方法

在静态图像人脸检测方法中,基于表象的方法是很重要的一个方法。总的来说,这种方法遵循一种统一的模式,即首先通过学习,在大量训练样本集的基础上建立一个能对人脸和非人脸样本进行正确识别的分类器,然后对被检测图像进行全局扫描,用分类器检测扫描到的图像窗口中是否包含人脸。若有,则给出人脸所在的位置。

采用这种检测模式的理论依据是:人脸具有统一的结构模式(都是由眉毛、眼睛、鼻子和嘴唇等人脸器官构成),如果把所有的图像集看作一个高维线性空间,那么整个人脸图像集仅对应于其中的某个子空间。于是可以通过检验待测图像窗口是否落在这个子空间中来判断其是否为人脸。因此,可以通过大量的人脸和非人脸样本来建立一个分类器,使它能够正确分辨这两种不同的图像模式,再利用训练好的分类器在未知的图像中检测人脸。采用这种检测策略的关键在于如何选取大量的具有代表性的图像样本,特别是非人脸图像样本,来训练分类器。

具体分类器的实现可以采用不同的策略,如采用神经网络的方法和传统的统计方法等。采用这种方法进行人脸检测的例子有 Sung 和 Poggio^[3]提出的基于样本学习的人脸检测方法, Turk, Pentland, Moghaddam 等^[4~6]提出的基于主成分分析(PCA)的人脸检测方法以及 Rowley 等^[7,8]实现的基于神经网络的方法等。

2.1 基于统计的人脸检测方法

Sung 和 Poggio^[3]提出了一种基于样本学习的人脸检测方法,可以在复杂背景中检测正面垂直人脸。他们将人脸检测问题归结为从图像样本中识别出人脸模板的学习过程,通过一个大约有1000幅包含人脸的图像数据库建立了人脸的分布

^{*}本文研究工作得到国家自然科学基金和江苏省自然科学基金资助。

模型,然后在一系列的人脸和非人脸样本中训练分类器得到区别人脸和非人脸模式的参数及阈值。他们所实现的系统包括两个主要部分:一个人脸分布模型和一个神经网络分类器。他们首先用 19×19 大小的样本图像建立标准的垂直正面人脸分布模型,并采集大量的人脸和非人脸样本训练神经网络分类器;为了能够判断新的图像窗口中是否包含人脸,首先将图像窗口正规化成 19×19 大小,然后计算被测窗口和已建立的人脸分布模型之间的差异度量,将结果提供给神经网络,根据其输出判断该窗口中是否包含人脸。

在 Sung 和 Poggio 的系统中,他们通过4150幅正面人脸图像来建立人脸的分布模型。所有的人脸样本首先进行预处理(放缩、光照修正、直方图均匀化、掩码等),然后通过聚类形成六个标准人脸模式类,每个人脸模式类是一个多维高斯分布函数,包括一个质心和一个用于描述质心周围局部数据分布的协方差矩阵。由于不可能得到所有的人脸样本,通过这种方法所建立的人脸模式类仅是对现实情况的模拟。同时,在现实世界中有很多的非人脸图像看起来像人脸,仅仅依靠六个人脸模式类还无法完全刻画人脸的分布模型,所以在 Sung 和 Poggio 的系统中,又通过同样的方法,在6189幅非人脸图像数据的基础上建立了六个非人脸模式类,Sung 和 Poggio 提出了一种 bootstrap 算法来获取具有代表性的非人脸图像集。通过这六个人脸模式类和非人脸模式类,可以建立起一个比较全面的人脸分布模型。

为了训练神经网络分类器,必须建立全面而又具有代表性的训练样本集。训练时,用于训练的样本与人脸分布模型中的十二个模式类进行比较,得到一个十二维的二值距离向量,将这个向量作为输入提供给神经网络分类器进行训练。分类器采用标准的 BP 学习算法。

在得到了人脸分布模型和神经网络分类器后,对于每幅被检测的图像,系统按照一定的窗口大小(19×19 到 100×100)扫描整个图像。当一张人脸被检测到后,系统就在相应的位置用方框标出。由于图像被缩放到不同的大小进行扫描检测,因此同一个位置可能有多个方框存在。Sung 等在两个数据集上测试了他们的人脸检测系统,第一个数据集包含71个不同人的301幅正面和接近正面的人脸图像,检测正确率为96.3%,其中有3个误识;第二个数据集包含149个人脸,共23幅图像,检测正确率为79.9%,其中有5个误识。

这种方法所存在的缺点主要是:只能在图像中检测垂直正面的人脸;用于建立人脸模型和训练神经网络的样本数量太大;进行全局搜索的时间较长。但是,其提出的通过建立标准人脸模型来检测人脸的思想启发了以后的研究。

在主成分分析(PCA)方法^[4]中,具有最大特征值的若干特征向量被提取出来构成一个子空间,这些特征向量作为子空间的基,它们保留了原数据中主要的信息而忽略了次要的部分。例如,从数据库中提取大量的人脸模式,经预处理后得到人脸的训练样本集。在这个人脸样本集中做主成分分析,得到一组特征向量,称为特征脸(Eigenface)。设 $\Sigma = \{\xi_1, \xi_2, \dots, \xi_n\}$ 为一组按特征值从大到小排序的特征脸,它构成了人脸特征空间的基, $f=f(x,y)$ 是一幅人脸图像。被测区域通过预处理后,可以表示为特征脸的线性表示。即:

$$\tilde{f} = x_1 \cdot \xi_1 + x_2 \cdot \xi_2 + \dots + x_n \cdot \xi_n \quad (1)$$

在进行人脸检测的过程中,被检测的图像区域根据公式(1)得到人脸子空间的表示了,由此可以计算与原图像的信噪比,即:

$$\beta = 10 \lg \frac{f}{f - \tilde{f}} \quad (2)$$

根据其是否大于某一阈值,可以判断被检测部分是否包含人脸。

Pentland 等^[5,6]提出了一种基于特征子脸的人脸及人脸特征的检测方法。其基本思想是这样的:人脸是一种器官的组合,如果将两幅人脸图像中对应的某一器官进行调换,它们分别作为人脸的属性不会受到影响,对于一幅正面人脸,可以根据人脸器官将其划分为若干区域;对于某一特定器官,存在一组基本器官图像(特征器官),该器官可以由这组特征器官的线性组合表示。由此,Pentland 建立了特征眼、特征唇等,通过计算被测图像区域与特征子空间的距离(DFFS)是否小于某一阈值来判断其是否属于该特征子空间。

1995年 Moghaddam 等^[7]指出,DFFS 仅是一个完整的目标概率函数的一部分,他们定义了一个最大似然检测策略(ML, Maximum Likelihood),其概率密度函数包含两个部分:DFFS(distance-from-feature-space)和 DIFS(distance-in-feature-space)。在他们做的实验比较中,ML 表现得比 DFFS 优越,ML 到达的最高的检测率是95%,在相同的检测率下,ML 比 DFFS 的误检测率要低^[9]。这主要是因为单一 DFFS 不能区分在特征空间中检测目标投影与特征空间原点的距离关系。

在文[10]中,Moghaddam 等为上述检测策略提出了完整的理论框架。他们总结了基于特征空间分解的目标检测方法,利用图像空间中的特征向量分解,建立了两类训练样本的概率密度估计。对于同尺度和方位的被测物样本,其分布假设为高斯分布,而其似然估计可表示为样本主空间概率分布与其正交补空间概率分布的乘积,或者为 DFFS 距离与 DIFS 距离之和;对于不同尺度和方位的被测物样本,由于其分布不再满足高斯分布,需要用混合高斯函数来表示 DIFS。基于此,他们建立了完整的最大似然检测策略,并将其应用于人脸检测、局部特征检测和手掌检测等领域。

2.2 基于神经网络的方法

在基于神经网络的人脸检测的方法中,值得一提的是由 Rowley 等^[7]人所所做的工作。他们的系统分为两个阶段。第一阶段是一个基于神经网络的分类器。它接受大小为 20×20 的图像窗口作为输入,产生一个从-1到1范围内的输出来标明该图像窗口中是否包含人脸,其中作为输入的图像窗口在提供给神经网络之前采用了 Sung 和 Poggio 的方法中同样的预处理过程。为了提高神经网络进行分类的精确度,需要大量的人脸和非人脸样本进行训练。他们大约采集了1050张人脸样本图像,这些图像中所包含的不同大小、方向、位置的人脸都必须经过正规化后提供给网络训练;而对于非人脸样本的采集,他们采用了 Sung 和 Poggio 提出的 bootstrap 算法,即首先建立一个小的非人脸样本集,和人脸样本集一起训练神经网络分类器,然后用训练好的分类器对不包含人脸的图像进行分类,将结果错误(被识别为人脸)的图像加入到原来的非人脸样本集合中,如此反复多次。第二阶段是合并重复检测位置并做出判断。由于在对输入图像进行扫描检测时,为了检测不同大小的人脸,对原图像做了放缩处理,因此对于一个正确的人脸位置,可能有若干个窗口标识它。为了给出一个唯一的检测结果,在第一阶段给出的所有检测位置中采用了与、或、投票等操作,得到一个最恰当的包含被检测人脸的窗口。

以上所提到的人脸检测方法,都是针对垂直正面人脸的,

Rowley 等^[8]又提出了一种改进方法,用于检测图像中按任意角度偏转的人脸。简单地把文[7]中的方法扩展用于检测按不同角度偏转的人脸,可以将整个输入图像不断地旋转一定的角度再进行检测。然而通过这样的改进所实现的系统非常耗时。例如,Rowley 等^[7]在1998年实现的人脸检测系统中,对于人脸图像有近似10度的旋转不变性(顺时针或逆时针),这就是说,对于一幅输入图像,要重复检测18次,每次将输入图像旋转20度。他们所提出的改进算法^[8]与前述方法的不同之处是增加了判断图像旋转角度的神经网络。整个检测系统由两部分构成:第一个部分是一个被称为 Router 的神经网络,用于人脸检测前对输入图像进行分析,它的输入是一个图像窗口,不论窗口中是否包含有人脸,将返回一个旋转角度,然后根据这个角度对该图像窗口进行反旋转、直方图均匀化等;第二个部分是进行人脸检测的神经网络,它的输入必须是经过预处理的图像窗口,根据结果输出可判断窗口中是否包含人脸,由于它的输入都是经过旋转后的垂直正面图像,每幅图像仅需判别一次就可得到结果,大大缩小了人脸检测所需的时间。值得注意的是第一部分中的神经网络并不负责检测人脸,如果输入的窗口图像中不包含人脸,则其输出将是一个毫无意义的旋转角度,原图像将照常被反旋转,然后提供给下一部分的神经网络进行检测。他们实现的系统在180幅有复杂背景的图像中能够检测到79.6%的人脸,其中包括一些误检测的位置。他们的工作与 Sung 等的不同之处在于:Rowley 等提出的神经网络直接用于判断图像窗口中是否包含人脸,其训练也需要大量的人脸和非人脸样本;而 Sung 实现的神经网络并不直接作用于原图像,而是把被测图像与人脸模型之间的距离向量作为网络的输入,其实现要比 Rowley 等的神经网络简单。

Feraud 等^[11]提出了一种新的基于神经网络的人脸检测方法。他们实现的系统可以同时处理灰度图像和彩色图像,并可用于检测侧面人脸。对于彩色图像要经过一个色彩过滤器,去除大约60%以上的背景。第一阶段,所有的图像经过预处理后由一个简单的神经网络进行初步检测,这一步去除99%以上的人脸区域假设。第二阶段是一个模块化的系统,它由四个新型的神经网络 CGM^[11]构成,每个 CGM 对应不同的人脸偏转角度,每幅被测图像在 CGM 输出层被重构,根据重构图像和原图像的差异(重构误差)来判断是否包含人脸。最后,Feraud 将他们的系统应用到了 Web 页面上的图像索引,结合文本信息,可以检测出特定的人脸图像。

3. 基于特征的人脸检测方法

基于特征的人脸检测方法是通过对不同的人脸面部特征的位置,然后根据它们之间的空间几何关系来定位人脸。这种方法又可进一步分为基于知识和基于局部特征的检测策略。基于知识的方法首先定位候选人脸区域,然后再通过人脸的先验知识来检验人脸是否存在。与之相对应的基于局部特征的方法中,人脸的局部特征如眼睛、鼻子和嘴唇等首先被检测出来,然后由这些局部特征组合成人脸。

3.1 基于知识的方法

这种方法是利用对人脸的先验知识导出的规则来进行人脸检测的。人脸局部特征的分布总是存在着一定的规律,例如人的两个眼睛总是对称分布在人脸的上半部分,鼻子和嘴唇中心点的连线基本与两眼之间的连线垂直等。于是,我们可以利用一组描述人脸的局部特征分布的规则来进行人脸检测。

这些规则可以表述为人脸局部特征之间的相对距离和位置关系等,当满足这些规则的图像区域找到后,则认为一幅人脸已被检测出来,然后可以对候选的人脸区域进行进一步的验证,以确定候选区域是否包含人脸。

Yang 和 Huang^[12]采用分层次的基于知识的方法进行人脸检测,他们所实现的系统共分三层,原始的数字图像作为系统的输入。在第一层,通过一些规则,在整幅图像的范围进行全局搜索,检测出所有可能的候选人脸区域。第二层,在上一层检测出的人脸候选区域的基础上进行进一步的检测,所用的规则基于更详细的面部特征。第三层,在上一层检测的人脸候选区域中进行直方图均匀化,然后进行眼睛和嘴唇的边界检测,如果被检测出的结果符合眼睛和嘴唇的特征,那么一张人脸也就被确定下来了。在上述第一、二层中,用于检测的图像进行了马赛克处理,马赛克图像通过降低原始图像的分辨率得到。

基于知识的方法所存在的问题主要是:识别率不高,如果图像背景中存在类人脸区域,则必然会导致误检测;该方法中的前两个步骤是一种由粗到细的检测策略,对于一个正确的人脸会被多次检测到,而后面的步骤中并没有给出如何确定最恰当的人脸位置的方法;对于不同视角的人脸,由于某些人脸特征不可见,所以不能使用这种方法检测;用于描述人脸特征之间关系的规则不易设计,规则制定得过高或者过低会造成拒识或误识。虽然这种方法存在着许多的局限性,但是它采用的对图像进行马赛克处理、分层次检测和通过先验知识指导人脸检测的思想对以后的研究提供了帮助。

3.2 基于局部特征的方法

与基于知识的方法相比,这种方法首先在整个图像中搜索一组人脸局部特征,然后通过它们之间的几何关系组合成候选的人脸区域。由于局部人脸特征的不变性,通过组合人脸局部特征,可以把这种方法应用在检测不同位置、角度的人脸。该方法所面临的主要问题是,由于图像噪声等因素的影响,造成人脸局部特征不明显,使得根据局部特征组合人脸的算法失效,从而产生拒识和误识。

Yow 和 Cipolla^[13]对基于局部特征的方法进行了改进,使它更适合于检测图像中不同大小、方位和视角的人脸。他们对原方法的改进主要体现在运用更多的人脸局部特征空间几何关系(image evidence)以及组合后的标准人脸模型。由于使用了大量的规则,满足单个规则的要求就降低了,这样使系统对于图像噪音有一定的适应性,不会产生过多错误的人脸候选区域。整个人脸检测过程是这样的:首先在原图像中找出若干影响点(interest point),它们可能是人脸局部特征的位置;在众多的影响点中进一步找出特征区域;根据预先定义的人脸模型,利用人脸局部特征的空间几何关系以及灰度值得到的局部特征组合成人脸,得到候选人脸区域的位置;最后在候选人脸区域中判定正确的人脸。在 Yow 和 Cipolla 实现的系统中,对包含110张不同人脸的图像进行了测试,正确率达到84.5%。这种方法的缺陷主要在于:不用于检测图像中小尺度的人脸,检测算法计算时间过长。

Han Chin-Chuan 等^[14]提出了一个基于形态学预处理的快速人脸检测方法。这个检测系统分为三个主要步骤:在第一个步骤中,一种基于形态学的技术用于眼睛近似物分割。在原图像中采用诸如闭运算、或运算等形态学操作定位眼睛近似物像素的位置,然后用标记过程产生眼睛近似物分段。在第二个步骤中,利用这些眼睛近似物分段来指导搜索潜在的人脸

区域。如果有类似眼睛、鼻子、眉毛和嘴唇的几何关系存在,那么一个临时的人脸区域就被标示出来。系统的最后一个步骤是人脸区域验证,首先在上一步中得到的所有候选人脸区域被正规化成 20×20 的图像,然后所有这些被正规化的候选人脸图像被输入到一个训练过的BP网络用于鉴别人脸,正确的人脸位置通过优化代价函数来确定。他们的方法可用于定位多角度的人脸位置,而且基于形态学的眼睛近似物分段过程去除了一幅复杂背景图像95%以上的区域,从而大大降低了下一步人脸检测所需的时间。这个算法在包含130个人脸的122幅图像中进行了测试,结果有122个人脸被准确地定位,显示了较高的检测率。在一幅 512×340 的图像中完成检测工作平均需要20秒。实际上,这个方法的一、二步采用了基于局部特征的方法,而最后的人脸验证过程则类似于Rowley等的基于神经网络的方法,综合了两种检测策略的长处,从而提高了检测速度。

Bur1^[15]等人建立了人脸局部特征之间相对位置关系的概率描述,能够识别经过平移、旋转过的人脸。他们的检测策略分为两个阶段:首先,用特征检测器搜索整个图像,找到组成人脸的各个部分,其中采用了基于弹性模板的软检测策略;最后,所有被检测出的局部特征组合成候选人脸,选择其中最符合人脸空间分布的作为检测出的人脸。这种方法与以前方法的不同之处在于采用了基于弹性模板的软检测策略,能克服图像光照变化等的影响。实际上,这种方法是基于局部特征和基于模板这两种检测方法的结合。

4. 基于模板的人脸检测方法

早期的基于模板匹配的方法^[16]是这样做的:首先建立一个标准的人脸模板,由包含局部人脸特征的子模板构成,然后对一幅输入图像进行全局搜索,对应不同尺度大小的图像窗口,计算与标准人脸模板中不同部分的相关系数,通过预先设置的阈值来判断该图像窗口中是否包含人脸。在文[16]中,Poggio利用基于模板的方法来定位眼睛的位置,他们把标准的眼睛模板调整为5个不同的尺寸,然后在输入图像中寻找眼睛。这种简单模板匹配的方法易于实现,但是也存在着缺点:图像噪声对检测结果影响很大,因此需要对输入图像做适当的预处理;标准模板的大小是人为设定的,因此不能动态检测眼睛的位置。

Yullie等人^[17]提出了基于弹性模板的方法,用于人脸特征的抽取。弹性模板^[18]是由一个根据被测物体形状而设定的参数化的可调模板和与之相应的能量函数所构成,能量函数要根据图像的灰度信息、被测物体轮廓等先验知识来设计。当用弹性模板进行人脸检测时,首先,将可调模板在待测图像中移动,并动态地调整其参数,计算能量函数。当能量函数到达最小值时,根据其位置和参数所决定的可调模板形状应该达到与人脸形状的最佳拟合,这样就检测到了一幅人脸。这种方法的优点在于,由于使用的弹性模板可调,能够检测不同大小、具有不同偏转角度的物体。但是其缺点是检测前必须根据待测物体的形状来设计弹性模板的轮廓,否则会影响收敛的结果;当对图像进行全局搜索时,由于要动态地调整参数和计算能量函数,计算时间过大。

结论 近二十年来,研究者在人脸检测方面取得了许多成绩,但是要设计并实现一个健壮的人脸检测系统,以下的一些问题还有待于解决:

• 图像自身的影响:光照条件将给建立一个健壮的人脸检

测系统带来很多影响。人脸实际上是三维线性空间中具有立体感的图像,在不同的光照条件,或者不同的面部表情下会得到不同的结果。

• 建立用于测试的标准数据库:目前存在很多的人脸检测方法,但是,为了对这些方法进行比较和评测,需要一个公正的测试环境,这就需要一个标准的人脸库。

• 多种检测方法的结合:目前的人脸检测方法都各自存在着缺点和不足,基于特征的方法较快速但是准确性低,基于表象的方法检测率高,但是训练和搜索的时间都比较长。考虑将几种检测策略结合在一起,将是一种有益的尝试。

除了本文提及的人脸检测方法之外,还有一些很有前途的研究方向,例如混合子空间方法^[19]和基于SVM的方法^[20]等。

人脸检测问题是一个艰巨而有趣的问题,目前所提出来的解决方法很多,但还没有一个统一的标准。它是实现全自动人脸识别系统的关键一步,是开发智能化的人机交互系统的关键技术。因其应用前景广泛且具有较大的难度,引起了研究者的广泛关注。本文仅对于静止图像中进行人脸检测的方法分别进行综述,希望能对今后的人脸检测研究工作有所帮助。

参考文献

- 1 Yang M-H, Ahuja N, Kriegman D. A survey on face detection methods. <http://vision.ai.uiuc.edu/mhyang/papers/survey.ps.gz>.
- 2 Yang M-H, Ahuja N. Detecting human faces in color images. In: Proc. of the Intl. Conf. on Image Processing, 1998. 127~130
- 3 Sung K-K, Poggio T. Example-based learning for view-based human face detection. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1998, 20(1): 39~51
- 4 Turk M, Pentland A. Eigenfaces for recognition. Journal of cognitive neuroscience, 1991, 3(1): 71~86
- 5 Pentland A, Moghaddam B, Starner T. View-based and modular eigenspaces for face recognition. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, Seattle, WA, 1994. 21~23
- 6 Moghaddam B, Pentland A. Face recognition using view-based and modular eigenspaces. In: Proc. of SPIE 2257, 1994. 12~21
- 7 Rowley H-A, Baluja S, Kanade T. Neural network-based face detection. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1998, 20(1): 23~38
- 8 Rowley H-A, Baluja S, Kanade T. Rotation invariant neural network-based face detection. In: Proc. of the IEEE Intl. Conf. on Computer Vision and Pattern Recognition, 1998. 38~44
- 9 Moghaddam B, Pentland A. A subspace method for maximum likelihood target detection. In: Proc. of the IEEE Intl. Conf. on Image Processing, 1995, 3: 512~515
- 10 Moghaddam B, Pentland A. Probabilistic visual learning for object representation. IEEE Trans. on pattern analysis and machine intelligence, 1997, 19(7): 696~710
- 11 F'eraud R, Bernier O, Viallet J E, Collobert M. A fast and accurate face detector for indexation of face images. In: Proc. of the 4th IEEE Intl. Conf. on Automatic Face and Gesture Recognition, 2000. 77~82
- 12 Yang G, Huang T-S. Human face detection in complex background. Pattern Recognition, 1994, 27(1): 53~63

(下转第110页)

这种方法使得开发和维护语法分析器非常简单。

(2) 基于 XML 的方法更加促进了代理技术在面向 WWW 中的应用。基于 XML 的内容很容易进行语义分析并且可以在 Web 浏览器中显示。

(3) 更进一步的使用是在 XML 中可以将链接(LINK)加入到 ACL 消息中^[9],这使得对 ACL 消息的内容共享和实例化很容易实现。ACL 中的参数值将不再只是 Token,而是可以指向对象或某些定义的链接 Links。例如:上面例子中 receiver 可以写作:

receiver:link="http://www.neu.edu.cn/people/abc"

也就是说它的参数可以是一个指向网络位置的链,通过这个链可以得到关于代理的一些详细的信息,如它的拥有者、联系信息、通信原语及用于接受通信消息的协议和端口号等等。这些信息是与另一个代理进行通信时所必需的扩充信息。

5. 基于 XML 的代理通信体系结构

按照上面的分析,可以得到集成 XML 的 KQML 的通信结构如图3所示。

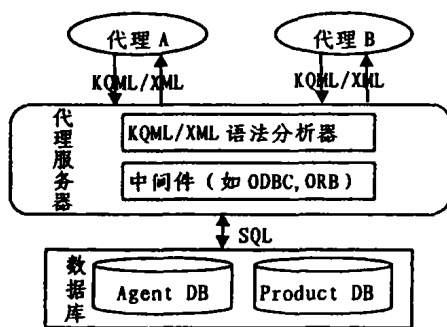


图3 通信组件

KQML 语法分析器和 XML 语法分析器是主要部件,对于一条基于 XML 的 KQML 动作表达式,KQML 语法分析器识别动作表达式的类型,提取其参数,XML 语法分析器根据 DTD 定义,编译动作表达式的内容层,然后触发相应的动作。代理通信过程可大致描述如下:

(1) 注册。所有的代理都是在代理服务器中运行的,为代理的运行提供相应的环境,所以代理必须在代理服务器中注册。对那些比较复杂的系统可能是在分布式平台上开发自己的应用,如 CORBA,DOCM 等。从而代理可作为一个 CORBA 对象而存在。代理的注册可以通过 CORBA 的 ORB 来完成。

(2) 编译动作表达式。请求代理(如 agentA)发出一个动

作表达式(如 ORDER),提交到代理服务器。代理服务器中的 KQML 语法分析器首先依据 KQML 规范,提取参数内容,并将内容层的信息交给 XML 语法分析器编译。

(3) 执行动作表达式。XML 语法分析器依据 DTD 的定义生成标准的 SQL 语句交给中间件(ODBC),中间件执行 SQL 语句,返回结果。

(4) 定位目标代理。KQML 语法分析器将提取的目标代理 ID 号交给中间件(ORB),中间件完成目标代理的定位。

(5) 完成请求。代理服务器将完成的结果发送给目标代理。

结论 本文讨论了基于 XML 的 KQML 通信技术的原理与实现。通过分析可知,XML 可以被很好地用于代理之间的通信与交互。在文[6]中,讨论了将 XML 和 KQML 结合在一起的代理通信机制在电子商务中的应用,并得到了较好的性能。XML 与 KQML 的结合可以用于基于其它通信机制实现代理通信的系统中,如基于 blackboard,基于 meeting^[8]的方式等的通信内容的表达也可以采用这种方式,同时这种方法也可以使用到基于其它通信语言实现通信的代理系统中,如 FIPA 中,因此这种方法具有广泛的应用价值。可以预见在将来的工业应用中,理想的 ACL 中的消息表达是基于 XML 的,而不是纯 ASCII 的^[9]。

参考文献

- 1 朱森良,邱瑜. 移动代理系统综述. 计算机研究与发展, 2001, 38 (1): 16~25
- 2 Finin T, Llabrous, Mayfield J. KQML as an agent communication language; in: Bradshaw, J.: software agents, Menlo Park 1995
- 3 Extensible Markup Language (XML). W3C Recommendation. <http://www.w3.org/XML/>
- 4 FIPA Agent Communication. <http://www.fipa.org/>, March 22, 2000
- 5 史忠植. 智能主体及其应用. 北京: 科学出版社, 2000
- 6 Karl K, Frank T, Daniel S. XML-based Agent Communication for Electronic Marketplaces
- 7 李薇, 李志署, 尹皓. XML-新的 Web 开发工具. 计算机应用, 1999, 19(5)
- 8 Domel P, Lingnau A, Drobniak O. Mobile Agent Interaction in Heterogeneous Environment. In: Proc. of the International Workshop on Mobile Agents, Lecture Notes in Computer Science, Springer-Verlag (D), No. 1219, April 1997. 136~148
- 9 Benjamin N G, Yannis L. An Approach to using XML and a Rule-based Content Language with an Agent Communication Language. In: proc. of the IJCAI-99 workshop on agent communication language, Held Stockholm, Sweden, 1999
- 17 Yullie A, Hallinan P, Cohen D. Feature extraction from faces using deformable templates. Intl. Journal of Computing Vision, 1992, 8 (2): 99~111
- 18 Jain A-K, Zhong Y, Dubuisson-Jolly M-P. Deformable template models: A review. Signal Processing, 1998, 71(2): 109~129
- 19 Yang M-H, Ahuja N, Kriegman D. Face detection using mixture of linear subspaces. In: Proc. of the 4th Intl. Conf. on Automatic Face and Gesture Recognition, 2000. 70~76
- 20 Terrillon J-C, Shirazi M N, Sadek M, Fukamachi H, Akamatsu S. Invariant face detection with support vector machines. In: Proc. of the 15th Intl. Conf. on Pattern Recognition, 2000, 4: 210~217

(上接第137页)

- 13 Yow K-C, Cipolla R. Feature-based human face detection. Image and Vision Computing, 1997, 15(9): 713~735
- 14 Han C-C, Liao H-Y, Yu G-J, Chen L-H. Fast face detection via morphology-based pre-processing. Pattern Recognition, 2000, 33 (10): 1701~1712
- 15 Burl M-C, Weber M, Perona P. A probabilistic approach to object recognition using local photometry and global geometry. In: Proc. of the European Conf. on Computer Vision, 1998. 628~641
- 16 Brunelli R, Poggio T. Face recognition: Features versus templates. IEEE Trans. on pattern analysis and machine intelligence, 1993, 15 (10): 1042~1052