

文章编号:1001-9081(2006)04-0776-04

建立 Avatar 感知能力与自主行为选择模型的新方法

陈中育,周晴红,朱响斌

(浙江师范大学 信息科学与工程学院,浙江 金华 321004)

(czy@zjnu.cn)

摘要:在人脸重构系统与智能虚拟环境系统的研究基础上,提出了一种基于 Agent 的建立虚拟人感知能力与行为选择模型的新方法。该方法把虚拟人(Avatar)当成一个 Agent,它在虚拟场景中能够识别三维物体以及利用传感器计算物体的运动方向。Avatar 的行走采用骨架模型来进行仿真,并且其行走行为由七个基本动作行为组成。实验结果显示新方法是具有较好的效果。

关键词:智能虚拟环境;行为动画;智能感知;行为决定

中图分类号: TP391.41;TP391.9 **文献标识码:** A

New perception and autonomous behavior modeling approach for Avatar

CHEN Zhong-yu, ZHOU Qing-hong, ZHU Xiang-bin

(Information Science and Engineering College, Zhejiang Normal University, Jinhua Zhejiang 321004, China)

Abstract: Based on the research of 3D facial model reconstruction and intelligent virtual entrainment, a new perception and autonomous behavior modeling approach for Avatar was proposed. In the new method, the Avatar was considered as an Agent, which could recognize 3D objects in the environment and calculate the direction of the moving objects by using the Avatar's perception sensor. Avatar walk behavior was simulated by skeleton model and denoted by seven motion sequences. Experiment results show that the new approach is more efficiency.

Key words: Intelligent Virtual Environment(IVE); behavior animation; intelligent perception; behavior decision

0 引言

在 Avatar 的行为表达中,主要涉及的有 Avatar 的表情与肢体行为。而虚拟感知能力主要涉及的是虚拟视觉的能力。所谓虚拟视觉主要是指 Avatar 具有识别三维模型和确定物体在虚拟场景中位置的能力。而在行为表达中,Avatar 能够通过人脸与肢体动作来向其他虚拟人表达其情绪的能力。Avatar 的行为基于骨架模型,它的行为动画采用七个基本的动作行为来表达,每个动作行为采用了有限状态机来实现。Avatar 的真实感表情采用了表情克隆技术,其人脸模型采用肌肉模型。

人脸建模的一个主要内容是研究如何对人脸进行变形,或者建立一个变形控制模型。因此人脸动画的一个关键的研究问题是如何实时地对人脸进行变形或构造一个人脸合成模型。Norman Badler 等在这个领域已经作出了颇具成效的工作^[3-7]。他们采用参数动作表达方法(Parameterized Action Representation, PAR)^[4,5]使 Agent 能够行动、规划,并使它的行动或多个 Agent 之间的交互行为更具有合理性。除了能描述虚拟人动作的语义,PAR 系统还使自主性 Agent 能产生未来的行为以及能够控制 Agent 个性、情绪以及情感的动画参数。文献[8]展示了人类行为的许多行为细节可以通过智能 Agent 来模拟。这些 Agent 能通过对其他虚拟人的行为施加影响,再从环境中的事件感知来评价它的情感重要性。

本文在原有的人脸重构系统^[9]与智能虚拟环境系统^[10,11]的基础上提出了一种基于 Agent 的建立虚拟人感知能力与行为选择模型的新方法。

1 Avatar 的虚拟感知器

Avatar 的感知包括了两个方面:识别和冲突检测。显然 Avatar 最需要知道的是周围有哪些物体,它们的位置、形状、颜色或其他有关信息,也就是其视觉能力。另一方面,还需要一些能对信息进行加工的传感器,以帮助 Avatar 进行精确判断,使虚拟人能够进行一些较精确的行为,如实时避碰等。

本文为 Avatar 设计的虚拟感知器以虚拟视觉为主,同时,针对虚拟人自主行为的需要,设计了相应的辅助传感器,进行信息的过滤和加工。

1.1 Avatar 的虚拟视觉

虚拟人视觉以 Avatar 的眼睛为视点,绘制虚拟环境,得到视觉图像和相应的 z-buffer(由于速度上的原因,在本文的系统中,虚拟视觉的窗口尺寸是 64×64)。

Avatar 的视觉图像不是用来绘制在计算机屏幕上的,而是用于虚拟人看到的物体的。因此其视觉图像就不需要纹理、光照的信息。但是颜色对于视觉图像来说是非常重要的,因为虚拟人的三维物体的识别主要采用了颜色信息。采用颜色信息进行物体识别时,需要对物体进行颜色编码,以便从绘制出的图像中检索物体。对物体进行颜色编码的方案有很多选择,本文采用了一种较为简便实用的方案,将虚拟环境中指向物体对象的指针直接作为物体的颜色来进行绘制。因而,从虚拟视觉的视觉图像中检索物体信息时,可以将检索到的颜色直接转为指向物体对象的视觉图像像素的三维坐标,通过以下方程式计算:

收稿日期:2005-12-06 基金项目:浙江省自然科学基金重点重大资助项目(Z603262)

作者简介:陈中育(1964-),男,浙江浦江人,副教授,主要研究方向:软件工程、计算机图形学;周晴红(1981-),男,江苏盐城人,硕士研究生,主要研究方向:计算机软件与理论;朱响斌(1971-),男,浙江义乌人,副教授,博士,主要研究方向:软件工程、计算机图形学、实时系统。

$$q_proj = \begin{bmatrix} -1.0 + 2 * i / WIDTH \\ -1.0 + 2 * j / HEIGHT \\ Z_SCREEN / Z_MAX \\ 1 \end{bmatrix}$$

$$q_world = 1 - proj * (MP)^{-1} \quad (1)$$

这里 (i, j) 是视觉图像的二维坐标, q_proj 是透视坐标, q_world 是世界坐标, $WIDTH$ 是屏幕的宽度, $HEIGHT$ 是屏幕的高度, P 是投影矩阵, M 是视点矩阵。

图 1 所示的是系统的虚拟环境与虚拟视觉图像。

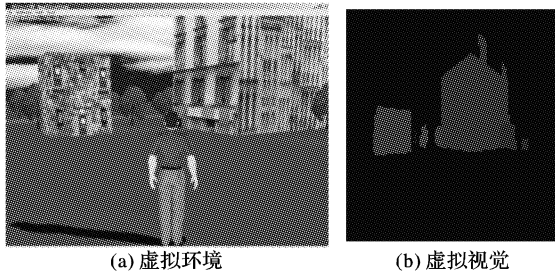


图 1 虚拟视觉图示

1.2 基于模型匹配的物体识别

虚拟人有了虚拟视觉以后,要以他的虚拟视觉识别三维物体是计算机视觉问题。在本文中采用了基于模型匹配的 Avatar 物体识别能力。所有的三维物体都可以用 $G = \{T, N, F\}$ 的集合来描述,其中 $T = \{Tr_1, Tr_2, \dots, Tr_k\}$ 是三角面片集合,每个三角面片由顶点 V_i 和法向量 N_i 来描述,因此 $Tr_i = \{V_i, N_i\}$ 。 N 表示三角面片间的相邻边,用 $N = \{e_{ij}\}$ 来描述。 F 是三角面片的特征向量函数,它是一个 s 维向量,因此 $F: Tr_i \rightarrow f_{mi} = (f_{i1}(Tr_1), f_{i2}(Tr_2), \dots, f_{is}(Tr_k))$,这里 $f_i(Tr_i)$ 是二维向量。

搜索匹配的三维模型是一个递归的过程。假设 G' 是需要识别的物体, G 是模型库中相应的模型,则匹配的过程如下:

1) 设一个匹配度 L , 其值在 G' 的 T' 与 G 的 T 之间。可以用集合来表示, $L = \{p_1, p_2, \dots, p_k\}$, 这里 $p_i = \{p_{i1}, p_{i2}, \dots, p_{ik}\} (i = 0, 1, 2, \dots, k)$ 表示了第 i 个三角面片的匹配度。

2) 定义匹配的优化函数如下:

$$E = \min \sum_{i=1}^k \sum_{j=1}^k ((Tr_i - M \cdot Tr_j') p_{ij})^2 + \sum_{i=1}^k \sum_{j=1}^k \sum_{o=1}^s (\|f_{oi} - f_{oj}'\| p_{ij})^2 \quad (2)$$

这里:

$$M = \lambda \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \quad (3)$$

M 是顶点变换矩阵, R 是平移向量, t 是旋转矩阵, λ 是放大因子。 M 可以通过手动对齐 G 与 G' 的 9 个点来得到。设 a 是阈值, 则 $E < a$, 就认为 G' 与 G 匹配。

1.3 虚拟传感器

传感器包括了目标传感器、障碍物传感器和对象传感器。这些传感器使 Avatar 实时地在虚拟场景中漫游。

目标传感器用来决定 Avatar 与目标之间的距离与角度。这个可根据如图 2 所示的方法,通过定位与方向来得到。Avatar 能通过目标传感器直接走向目标。

障碍传感器为 Avatar 进行冲突检测,且仅用于 Avatar 行走中的冲突检测。图 3 说明了障碍传感器工作原理,传感器的窗口由 (θ_{max}, d_{max}) 所定义,其中 θ_{max} 是相对于当前 Avatar 朝向的最大转角, d_{max} 是感知的最大的距离。障碍传感器返

回探测到的障碍物集合,通过计算出障碍物集合的力场,虚拟人可以选择出能够避开障碍物的方向。

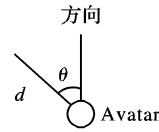


图 2 目标传感器

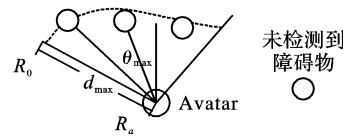


图 3 障碍传感器

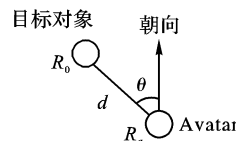


图 4 对象传感器

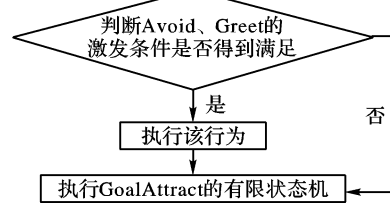
对象传感器与目标传感器的功能相似,只不过对象传感器返回的是虚拟人与目标对象的距离与朝向角度。为了防止可能出现的穿越现象,目标对象的大小以包围盒或包围圆柱来估算。在计算虚拟人与目标对象的距离时,选择以包围盒或包围圆柱来计算需要根据目标对象的特性来决定。图 4 中, d 是虚拟人与目标对象之间的距离, θ 是虚拟人的朝向与 R_o, R_e 之间的夹角。

目标对象传感器通常在 GoalAttract 行为中使用,使虚拟人朝目标对象行走。

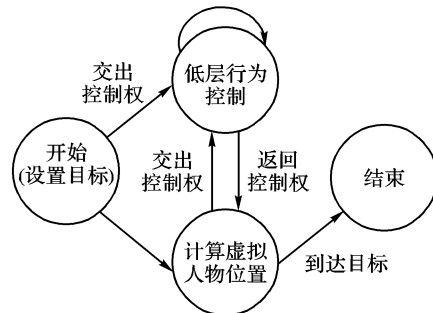
2 Avatar 自主行为的设计

自主行为模型的目标是在动态的、不可预测的虚拟环境中正确指导 Avatar 的动作,即行为模型要完成的任务是:在相互冲突的目标之间权衡,根据当时的环境选择出对虚拟人最适合的动作。行为模型一般提供一个可供执行的行为集合,以支持虚拟人在虚拟环境中的自主行为。

建立一个完整自主行为模型大致需要以下三个步骤: 1) 为 Avatar 选择适当的行为集合,要视 Avatar 在虚拟世界中担任的角色及 Avatar 要完成的任务而定,所选择的行为集合应能满足 Avatar 完成任务的需要; 2) 对行为分类以建立层次结构,能够有利于系统的构造,使 Avatar 在任一时刻能专注于某一种行为; 3) 设计行为选择机制,能解决如何从相互竞争的行为中挑选获胜行为的问题。



(a) 流程



(b) GoalAttract 行为的状态机

图 5 处理流程和 GA 的 FSM

2.1 行为层次结构

在行为集合中,行走行为是最主要的研究内容。Avatar 的行走仿真采用了运动链技术。在虚拟人的骨架中,大腿根

是运动链的根,而脚部是运动链的末端效应器。

Avatar 的行走可以用七个基本动作来描述: $WM = \{GE, AV, GA, PF, GL, GO, CH\}$ 。其中:

AV (Avoid) 表示避让;

GE (Greet) 表示打招呼,遇见熟人时驱动物体做出打招呼的动作;

GA (GoalAttract) 表示行走到一个可以直接达到的目标地点;

PF (PathFollowing) 表示沿路径行走;

GL (GoLocation) 表示到达指定地点;

GO (GoObject) 表示到达指定对象;

CG (Chasing) 表示跟踪指定对象。

这里面最重要的动作是 GA 与 PF,因为其他的五个动作都可以由这两个动作实现。GA 采用有限状态机(FSM)^[12]来描述。图 5 说明了 FSM 与 GA 的处理流程。PF 采用路径动态编程,它引导 GA 找到最终的目标。

2.2 行为集合与行为的生成

Avatar 的行为表达包括脸部表情和身体姿态两部分。在本文的动态虚拟环境中,Avatar 能表现他们脸部表情;同时 Avatar 还能通过身体姿态来表达他们的情感。首先,Avatar 行为与表情之间的有关定义如下:

定义 1 对于某一虚拟环境,行为集合 E 是有可能行为的并集。 $E = \{e_1, \dots, e_n\}$, n 是行为状态数。Avatar 行为是 E 集合中的一个元素。对于 E 集合中的两个行为 $E(A1)$ 与 $E(A2)$,它们的并用 $E(A1) + E(A2)$ 来表示,它们的交用 $E(A1) * E(A2)$ 来表示。

定义 2 对于某一虚拟环境,情感集合 P 是有可能姿态行为和相应人脸表情的集合。 $P = \{p_1, \dots, p_m\}$, m 是姿态与表情的状态个数, $P_j(j = 1, \dots, m)$ 是某一姿态或表情。

定义 3 对于某一虚拟环境,行为集合是 E ,情感集合是 P 。则设一个从 E 到 P 映射函数 fm 。对于任何 $x \in E, y \in P$,有 $y = fm(x)$,称 fm 为行为函数。

一般来说,行为函数采用产生式规则来表示,在本文的系统中,行为函数是产生式规则的集合,如下所示:

```

如果 Avatar 是注视的,则
{ 姿态运动:= 注视的动作
  //从运动数据库中选择合适的运动数据
  脸部表情:= 注视的表情 //从表情数据库选择合适的表情
}
否则 如果 Avatar 是害怕的 则
{ ... }

```

定义 4 对于某一虚拟环境,设 $E(A)$ 为行为集合 E 的一个子集, $P(A)$ 为 P 的一个元素。 fm 是从 E 到 P 的行为函数。如果对于任何 $x \in E(A)$,有 $P(A) = fm(x)$,那么 $E(A)$ 称为与 $P(A)$ 是关联的,可以写成 $E(A) \rightarrow P(A)$; 否则写成 $E(A) \mapsto P(A)$ 。

定义 5 行为源就是能引起 Avatar 情感行为的外部刺激。对于一个行为集合 E ,行为源可用称为源刺激的有限集合 $O(E)$ 来表示。

对于任一元素 $O(I) \in O(E)$,其中 $I \in [1, N], N$ 是 $O(E)$ 集合的元素个数。 $O(I)$ 的空间分布采用 $\Omega(I)$ 来表示。对于一个处于点 P 的 Avatar,如果 $\Omega(I) \mapsto P(x, y, z)$,则源刺激 $O(I)$ 将不会影响 Avatar 的行为。一般来说,对于每个 $O(I)$,让 $\Theta(O(I))$ 表示源刺激的强度,其中 $\Theta \in [0, 1]$ 。由于 $O(I)$ 的强度正比于 $\Theta(O(I))$,因此 $\Omega(I)$ 的中心是 $\Theta(O(I))$ 最大的

地方。

对于行为集合 E ,其中任一元素 e 是用于表达行为状态的描述语,譬如“高兴”。Avatar 抵抗行为 e 的强度用 $C(e)$ 来表示, $C(e) \in [0, 1]$ 。 $C(e)$ 的大小反比于 Avatar 的行为数。对于所有的 $O(I)$,如果 $C(e) - \Theta(O(I)) > 0$,那么 Avatar 对于行为源 $O(e)$ 不会产生行为 e 。如果存在 $O(I)$ 且 $C(e) - \Theta(O(I)) < 0$,那么 Avatar 对于行为源 $O(e)$ 将会产生行为。

设 $S(I)$ 表示 $P(x, y, z)$ 到 $\Omega(I)$ 中心的距离, t 为时间变量。一般来说,会有两种情况:一种是 $(S(I))/dt < 0$ 以及对应于 $\Omega(I)$ 的 Avatar 方法;另一种是 $d(S(I))/dt > 0$ 以及对应于 $\Omega(I)$ 的 Avatar 躲避方法。在虚拟环境中,如果行为源 $O(I)$ 是活跃的,那么开始时间与终止时间分别用 $Tm1$ 和 $Tm2$ 来表示。行为将延续 $(Tm2 - Tm1)$ 时间。

有限状态机用于表达行为的变化。一个行为状态可以用一个三元组 $\langle P(t), E(t), B(t) \rangle$ 来表示,其中 t 是时间变量, $P(t)$ 是从感知中得到的输入集,包括行为源与其他环境信息, $E(t)$ 是在时间 t 所有可能行为状态的行为状态, $E(t0)$ 表示行为源状态, $B(t)$ 为输出行为集。这里有两个函数: $L1: E(t) * P(t) \rightarrow E(t), L2: E(t) * P(t) \rightarrow B(t)$ 。 $L1$ 称为行为变化函数, $L2$ 称为行为输出函数。 $L1$ 和 $L2$ 可以用产生规则构成。在本文的系统中,虚拟世界生成采用世界地图和原始信息,原始信息包括了行为原始位置,Avatar 原始位置及行走速度。Avatar 的行为与表情由以下三步来生成:

- (1) 读取地图,在虚拟世界中选择导航点,引导角色沿着导航点移动,其中包括了一些源刺激。
- (2) 获得当前所有有效行为源,利用虚拟传感器与表情预处理机制选择最合适的表情和行为。
- (3) 在 Avatar 识别虚拟世界后,产生合适的情感表达

3 实验结果

实验系统由三个阶段构成:第一阶段是 Avatar 在虚拟环境中的识别能力,第二阶段是表情与行走行为的预测与决定,第三阶段是 Avatar 的表情克隆与行走。

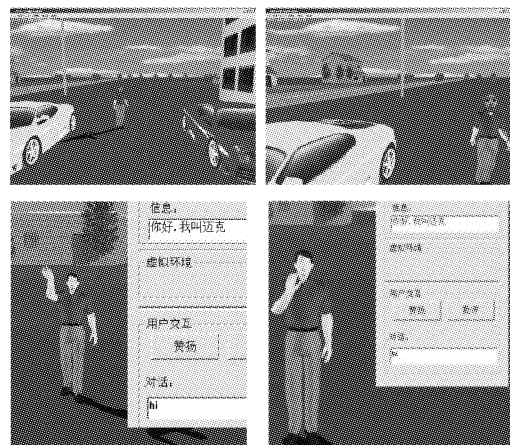


图 6 一些实验结果

Avatar 在虚拟世界中的行走,可能会碰到一些三维模型,因此需要识别三维模型。识别与冲突检测利用了虚拟传感器与 Avatar 的视觉。行走行为的实现采用了骨架模型并且其行为可以七个基本的动作来完成。表情克隆采用 RBF 对齐技术和人脸肌肉模型的变形来实现,因此 Avatar 能自然地表达他的表情。

以上所介绍的方法已经在我们的系统 IVES 中实现。实

验原型系统是在 Windows 2000 下用 VC++ 6.0 实现的,实验结果如图 6 所示。在该系统中,虚拟人的头是采用我们的方法建模的,虚拟人的身体是用软件 Poser 来完成的。图 6 说明了 A 先生遇到汽车时的脸部表情以及远近距离打招呼的行为表达。

4 结语

本文主要介绍了感知与 Avatar 在 IVE 中自主行为。Avatar 表情通过表情克隆与基于脸部肌肉模型的三维人脸重构成来仿真。这种方法同样可以应用于人机交互的其他领域。例如,如果对虚拟世界和 IF-ACTION 规则加以修改,再增加语音识别能力,Avatar 就能有语音驱动的脸部表情。

参考文献:

[1] EKMAN P, HUANG TS, SEJNOWSKI TJ, *et al.* Final report to NSF of the planning workshop on facial expression understanding, Human Interaction Laboratory[D]. University of California, San Francisco, 1993.

[2] PARKE FI, WATERS K. Computer Facial Animation[M]. AKPeters, Wellesley, Massachusetts, 1996.

[3] BINDIGANAVALA R, SCHULER W, ALLBECK JM, *et al.* Dynamically Altering Agent Behavior Using Natural Language Instructions [Z]. Autonomous Agents, 2000. 293 - 300.

[4] BADLER N, ALLBECK J, ZHAO L, *et al.* Representing and Parameterizing Agent Behavior[A]. Proceedings of Computer Animation [C]. Geneva, Switzerland: IEEE Computer Society, 2002.

[5] ALLBECK J, KIPPER K, ADAMS C, *et al.* ACUMEN: Amplifying Control and Understanding of Multiple Entities[A]. Proceedings of Autonomous Agents and Multi-Agent Systems[C]. Bologna, Italy: ACM Press, 2002. 191 - 198.

[6] ALLBECK J, BADLER N. Toward Representing Agent Behaviors Modified by Personality and Emotion[EB/OL]. <http://www.seas.upenn.edu/~allbeck/eca02.pdf>, 2005.

[7] BADLER N, PHILLIPS C, WEBBER B. Simulating Humans: Computer Graphics Animation and Control[M]. Oxford University Press, New York, 1993.

[8] GRATCH J, MARSELLA S. Tears and Fears: Modeling emotions and emotional behaviors in synthetic Agents[A]. Proceedings of the 5th International Conference on Autonomous Agents[C]. Montreal, Canada, 2001.

[9] LIANG R, CHEN C, PAN ZG, *et al.* A new algorithm for 3D facial model reconstruction and its application in VR[A]. Proceedings of SPIE[C]. 2003. 79 - 86.

[10] LIU Z, PAN ZG, XU WW. A method of Emotional Behavior Animation of Virtual Human, Workshop[A]. Proceedings of Virtual Environment on PC Cluster[C]. 2002. 277 - 283.

[11] LIU Z, PAN ZG, ZHANG MM. Behavior Animation of Virtual Human in Virtual Society[A]. The 12th International Conference on Artificial Reality and Telexistence(ICAT)[C]. 2002. 186 - 187.

[12] BADLER N, WEBBER B, BECKETT W, *et al.* Planning for animation[M]. Interactive computer animation, Prentice-Hall Press, 1996.

(上接第 775 页)

4 嵌入式监控系统的运行实例

本文实现了二次纤维造纸废水处理监控系统。图 6 是系统结构图,其中嵌入式设备采用 AMD 公司的 PCM-3350 主板。监控系统中还使用研华 ADAM4017 +、ADAM4024、ADAM4060 以及 ADAM4520 智能通讯模块连接嵌入式设备和法国 AWA 在线 COD 检测仪等设备。ADAM4017 + 是 8 路模拟量输入模块,用于获取 COD 仪的数据;ADAM4024 是 4 路模拟量输出模块,作用是控制加药流量;ADAM4060 是 4 路继电器输出模块,用于控制各个电磁阀的关闭;ADAM4520 是隔离 RS232 到 RS422/485 转换器,用于 ADAM4017 +、ADAM4024、ADAM4060 与嵌入式设备间串口协议转换。

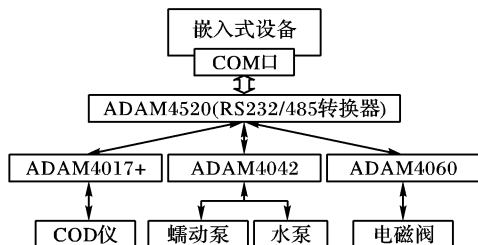


图 6 嵌入式监控系统结构体系

图 7 是监控系统运行的主界面,软件提供了人机界面、数据列表和实时趋势图的显示,同时还提供改变采集量标签和历史数据查询的功能。人机界面在线实时现场生产数据,从数据列表中可直观地看到一段时间内各个采集变量的变化情况。在历史数据查询中,可以查询不同时间段的监控变量数据变化情况,并将结果显示在数据列表中,方便用户了解生产过程。当组态工程需要变化时,采集量定义模块特别有用,使用它增加或者删除采集量标签时不必完全重新生成应用程序。

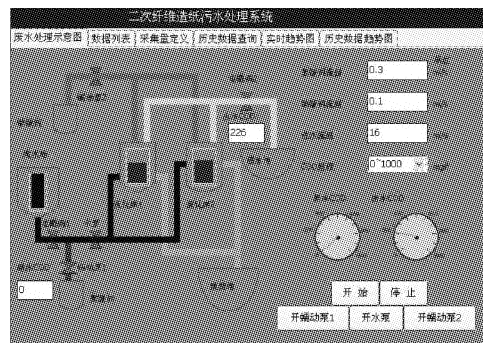


图 7 软件运行主界面

5 结语

本文介绍了基于 Windows CE.NET 的嵌入式监控系统的设计和实现技术,并以二次纤维造纸废水处理过程的监控加以验证。文中所构造的功能模块可实现一定程度的可重用,为进一步构造监控组态软件提供基础。程序的功能还需要完善,如上述工作只实现对系统的初步监控功能,要实现对系统的精确控制,实现对出口水质的稳定达标排放,还需要进一步完善控制算法模块等,该工作正在进行当中。

参考文献:

[1] 马国华. 监控组态软件及其应用[M]. 北京: 清华大学出版社, 2001.

[2] 毛峰, 马中. 面向嵌入式系统组态工具的研究[J]. 计算机与数字工程, 2003, 31(2): 12 - 14.

[3] 高海波, 马中. 面向嵌入式系统的组态支撑技术[J]. 舰船电子工程, 2004, 24(4): 56 - 58.

[4] 刘松青, 任志峰. 工业循环水动态模拟装置的计算机控制[J]. 仪表·自动化, 2003, 10(2): 57 - 59.

[5] 黄伟, 叶朝晖. 面向对象的工控组态软件的开发与应用[J]. 工业控制计算机, 1996, (4): 3 - 6.