



(12)发明专利申请

(10)申请公布号 CN 108877807 A

(43)申请公布日 2018. 11. 23

(21)申请号 201810726777.4

(22)申请日 2018.07.04

(71)申请人 广东猪兼强互联网科技有限公司
地址 510000 广东省广州市天河区元岗路
200号第九栋9213-9215房

(72)发明人 陈志林

(74)专利代理机构 北京华识知识产权代理有限公司 11530

代理人 刘艳玲

(51) Int. Cl.

G10L 15/26(2006.01)

G06Q 30/02(2012.01)

B25J 11/00(2006.01)

权利要求书1页 说明书4页 附图1页

(54)发明名称

一种用于电话销售的智能机器人

(57)摘要

本发明公开了一种用于电话销售的智能机器人,该智能机器人包括:语音采集模块、语音处理模块、语音识别模块和语音播报模块。语音采集模块,用于采集客户的语音信息;语音处理模块,用于对采集的语音信息进行处理,得到第一语音信号;语音识别模块,用于根据第一语音信号,确认客户的身份信息以及客户的业务需求;语音播报模块,用于根据客户的业务需求,向客户推送相应的业务信息。本发明提供的智能机器人可以替代电话销售人员为客户提供服务,避免了以往电话销售人员的进行电话推销时的态度不友善以及没有针对性地为客户提供服务的缺点,极大地提高了电话销售的成功率,而且节约了人力资源,具有推广应用价值。



1. 一种用于电话销售的智能机器人,其特征在于,包括:语音采集模块、语音处理模块、语音识别模块和语音播报模块;

所述语音采集模块,用于采集客户的语音信息;

所述语音处理模块,用于对采集的语音信息进行处理,得到第一语音信号;

所述语音识别模块,用于根据所述第一语音信号,确认客户的身份信息以及客户的业务需求;

所述语音播报模块,用于根据所述客户的业务需求,向客户推送相应的业务信息。

2. 根据权利要求1所述的智能机器人,其特征在于,还包括数据库,所述数据库用于存储各种业务信息的特征参数信息和客户的身份信息。

3. 根据权利要求2所述的智能机器人,其特征在于,所述语音识别模块包括客户身份识别单元和语音内容识别单元;

所述客户身份识别单元,用于根据所述第一语音信号,提取客户的声纹信息,确认客户的身份信息,并将确认好的客户的身份信息存储至所述数据库;

所述语音内容识别单元,用于根据所述第一语音信号,提取能够描述客户业务需求的特征参数,并确定所述第一语音信号中客户业务需求的特征参数与所述数据库中预存的业务信息的特征参数的匹配程度,当两者的匹配程度大于预设定的阈值时,通过所述语音播报模块向客户推送相应的业务信息。

4. 根据权利要求3所述的智能机器人,其特征在于,所述语音处理模块包括语音增强单元和语音去噪单元;

所述语音增强单元,用于对采集的语音信号进行语音增强处理;

所述语音去噪单元,用于对增强处理后的语音信号进行去噪处理,得到第一语音信号。

5. 根据权利要求4所述的智能机器人,其特征在于,所述对采集的语音信号进行语音增强处理,包括:

(1) 采用傅里叶变换将采集的语音信号从时域变换到频域内;

(2) 根据自定义的语音增强公式对频域内的语音信号进行语音增强处理,得到增强后的语音信号,其中,自定义的语音增强公式为:

$$\tilde{Y}(t, d) = \left[\frac{\exp[-D(t, d)]}{1 + \exp[-D(t, d)]} \times G_{th}^{p(t, d)} + G_{min}^{[1-p(t, d)]} \right] \cdot Y(t, d)$$

式中, $\tilde{Y}(t, d)$ 表示经增强处理后的语音信号, $D(t, d)$ 指的是先验信噪比, $Y(t, d)$ 表示增强处理前的频域内的语音信号, $p(t, d)$ 表示第 t 帧第 d 频带处语音存在的概率值, G_{min} 表示在非语音帧的状态下的最小增益值, t 表示帧数, d 表示频率, G_{th} 是设定的在语音帧的状态下的增益值;

(3) 采用傅里叶逆变换将增强后的语音信号从频域变换到时域,得到在时域内的增强后的语音信号。

一种用于电话销售的智能机器人

技术领域

[0001] 本发明涉及智能机器人技术领域,具体为一种用于电话销售的智能机器人。

背景技术

[0002] 传统的电话销售方法需要销售人员为客户推销商品,这种方法不仅需要投入大量的资金、设备,而且需要配备较多的销售人员,还要进行较长时间的人员培训,同时服务质量也深受销售人员的素质以及业务量大小的制约。

发明内容

[0003] 针对上述问题,本发明旨在提供一种用于电话销售的智能机器人。

[0004] 本发明的目的采用以下技术方案来实现:

[0005] 一种用于电话销售的智能机器人,该智能机器人包括:语音采集模块、语音处理模块、语音识别模块和语音播报模块。

[0006] 语音采集模块,用于采集客户的语音信息;语音处理模块,用于对采集的语音信息进行处理,得到第一语音信号;语音识别模块,用于根据第一语音信号,确认客户的身份信息以及客户的业务需求;语音播报模块,用于根据客户的业务需求,向客户推送相应的业务信息。

[0007] 有益效果:本发明提供了一种用于电话销售的智能机器人,该智能机器人可以替代电话销售人员为客户服务,避免了以往电话销售人员的进行电话推销时的态度不友善以及没有针对性地为客户提供服务的缺点,极大地提高了电话销售的成功率,而且节约了人力资源,具有推广应用价值。

附图说明

[0008] 利用附图对本发明作进一步说明,但附图中的实施例不构成对本发明的任何限制,对于本领域的普通技术人员,在不付出创造性劳动的前提下,还可以根据以下附图获得其它的附图。

[0009] 图1为本发明的框架结构图;

[0010] 图2为本发明语音处理模块的框架结构图;

[0011] 图3为本发明语音识别模块的框架结构图。

[0012] 附图标记:

[0013] 语音采集模块1;语音处理模块2;语音识别模块3;语音播报模块4;数据库5;客户身份识别单元31;语音内容识别单元32;语音增强单元21;语音去噪单元22;预处理子单元221;检测子单元222;更新子单元223。

具体实施方式

[0014] 结合以下应用场景对本发明作进一步描述。

[0015] 图1示出了一种用于电话销售的智能机器人,该智能机器人包括语音采集模块1、语音处理模块2、语音识别模块3和语音播报模块4。

[0016] 语音采集模块1,用于采集客户的语音信息;语音处理模块2,用于对采集的语音信息进行处理,得到第一语音信号;语音识别模块3,用于根据第一语音信号,确认客户的身份信息以及客户的业务需求;语音播报模块4,用于根据客户的业务需求,向客户推送相应的业务信息。

[0017] 有益效果:本发明提供了一种用于电话销售的智能机器人,该智能机器人可以替代电话销售人员为客户提供服务,避免了以往电话销售人员的进行电话推销时的态度不友善以及没有针对性地为客户提供服务的缺点,极大地提高了电话销售的成功率,而且节约了人力资源,具有推广应用价值。

[0018] 优选地,该智能机器人还包括数据库5,数据库5用于存储各种业务信息和客户的身份信息。

[0019] 优选地,参见图2,语音识别模块3包括客户身份识别单元31和语音内容识别单元32,客户身份识别单元31,用于根据第一语音信号,提取客户的声纹信息,确认客户的身份信息,并将其客户的身份信息存储至数据库5;语音内容识别单元32,用于根据所述第一语音信号,提取能够描述客户业务需求的特征参数,并确定所述第一语音信号中客户业务需求的特征参数与所述数据库5中预存的业务信息的特征参数的匹配程度,当两者的匹配程度大于预设定的阈值时,通过所述语音播报模块4向客户推送相应的业务信息,具体而言,是向客户推送匹配程度大于预设定的阈值的业务信息,本实施例中的业务信息可以是商品或者服务。

[0020] 有益效果:在上述实施例中,分别设置客户身份识别单元31和语音内容识别单元32,用于对客户的身分进行确认,同时将确认好的客户身份信息存储至数据库中,该做法有利于公司掌握住客户的身份信息以及客户的需求信息,有利于维系公司与客户之间的合作关系,方便公司为客户推送与之相适应的业务内容。

[0021] 优选地,参见图3,语音处理模块2包括语音增强单元21和语音去噪单元22,语音增强单元21用于对采集的语音信号进行语音增强处理,语音去噪单元22用于对增强处理后的语音信号进行去噪处理,得到第一语音信号。

[0022] 优选地,对采集的语音信号进行语音增强处理,包括:

[0023] (1) 采用傅里叶变换将采集的语音信号从时域变换到频域内;

[0024] (2) 根据自定义的语音增强公式对频域内的语音信号进行语音增强处理,得到增强后的语音信号,其中,自定义的语音增强公式为:

$$[0025] \quad \tilde{Y}(t, d) = \left[\frac{\exp[-D(t, d)]}{1 + \exp[-D(t, d)]} \times G_{th}^{p(t, d)} + G_{min}^{[1-p(t, d)]} \right] \cdot Y(t, d)$$

[0026] 式中, $\tilde{Y}(t, d)$ 表示经增强处理后的语音信号, $D(t, d)$ 指的是先验信噪比, $Y(t, d)$ 表示增强处理前的频域内的语音信号, $p(t, d)$ 表示第 t 帧第 d 频带处语音存在的概率值, G_{min} 表示在非语音帧的状态下的最小增益值, t 表示帧数, d 表示频率, G_{th} 是设定的在语音帧的状态下的增益值;

[0027] (3) 采用傅里叶逆变换将增强后的语音信号从频域变换到时域,得到在时域内的增强后的语音信号。

[0028] 有益效果:利用傅里叶变换将获取的语音信号变换到频域内并在频域内对语音信号进行增强处理,其中,在自定义的语音增强函数,通过引入 $p(t,d)$,使得在非语音帧的时候,也考虑估计干净语音的幅度,同时引入了 G_{min} 也进一步减少了语音失真,同时将残差噪声转变为白噪声,让人听上去更平滑,更舒适;该算法可以更好地保留语音信号的主要信息,去除低频部分的背景噪声。在利用语音增强函数对语音信号进行增强处理时,引入 G_{th} 进一步修正在语音帧内的语音信号,该做法能够进一步对有效语音信号进一步增强处理,有利于后续对有效语音信号中特征参数的提取。

[0029] 优选地,为进一步得到更加纯净的语音信号,利用下式对自定义的语音增强函数中的先验信噪比 $D(t,d)$ 进行修正,并将修正后的 $D(t,d)$ 代替语音增强函数中的 $D(t,d)$,其中,修正先验信噪比的计算公式为:

$$[0030] \quad \tilde{D}(t,d) = \frac{D(t,d)}{1 + D(t,d)} \times \left(1 + \frac{2 \times D(t,d) \times \vartheta(t,d)}{\vartheta^2(t,d) + D^2(t,d)} \right)^v$$

[0031] 上式中, $\tilde{D}(t,d)$ 表示修正后的先验信噪比, $\vartheta(t,d)$ 表示后验信噪比, v 为设定的调节因子,用于对噪声进行平衡控制。

[0032] 在对语音信号进行增强处理时,通过引入修正后的先验信噪比 $\tilde{D}(t,d)$ 来替换先验信噪比 $D(t,d)$,该做法能够进一步减小误差,提高了在低信噪比场景下进行语音增强处理时的增强效果,同时在对先验信噪比 $\tilde{D}(t,d)$ 进行修正时,噪声容易出现过估和欠估的情况,通过引入调节因子 v (通常 v 选取0.75),在后续对噪声功率谱进行估算时,能够有效防止过估和欠估的情况的发生,使得语音失真最小。

[0033] 优选地,语音去噪单元22包括预处理子单元221、检测子单元222和更新子单元223。

[0034] 预处理子单元221利用声学滤波器组将增强后的语音信号在时域分解为多个子带的信号;检测子单元222用于对每个子带信号进行能量检测,估计有效音频存在的区域;更新子单元223用于根据检测子单元222的检测结果,对当前帧的噪声能量估计值进行更新,并将更新后的噪声能量估计值应用于检测子单元222中,进而实现对有效音频存在的区域的更新。

[0035] 在一个实施例中,所述对每个子带信号进行能量检测,包括:

[0036] (1) 对各个子带的信号进行分帧处理;

[0037] (2) 利用能量计算公式计算每个子带信号经分帧处理后的帧信号的能量估计值和初始噪声能量估计值,其中, j 子带的能量计算公式为:

$$[0038] \quad E_{fr}(i,j) = \sum_{k=1}^{Len_{fr}} S(k,i,j)^2$$

$$[0039] \quad E_{noise}(i,j) = \frac{1}{m} \sum_{i=1}^m E_{fr}(i,j)$$

[0040] 式中, $S(k,i,j)^2$ 是 j 子带第 i 帧中的第 k 个采样点的幅值, Len_{fr} 为一帧信号的长度, $E_{fr}(i,j)$ 是第 i 帧信号的能量估计值, $E_{noise}(i,j)$ 是 j 子带第 i 帧信号的初始噪声能量估计值; m 是子带 j 的前 m 帧信号;

[0041] (2) 根据得到能量估计值和初始噪声能量估计值,利用判定公式对当前帧是否是有效音频进行判定,其中,判定公式为:

$$[0042] \quad V_{cross}(i, j) = \begin{cases} 0 & E_{fr}(i, j) < \eta_1 E_{noise}(i, j) \\ 1 & E_{fr}(i, j) > \eta_2 E_{noise}(i, j) \\ V_{cross}(i-1, j) & \text{else} \end{cases}$$

[0043] 式中, η_1 和 η_2 为阈值加权系数,用于调节阈值大小,且满足 $\eta_1 < \eta_2$; 当 $E_{fr}(i, j) < \eta_1 E_{noise}(i, j)$, $V_{cross}(i, j) = 0$, 此时第 i 帧为噪声帧, 当 $E_{fr}(i, j) > \eta_2 E_{noise}(i, j)$ 时, $V_{cross}(i, j) = 1$, 此时第 i 帧为语音帧, 在其他的情况下, $V_{cross}(i, j)$ 与上一帧的判别结果一致;

[0044] (3) 遍历所有子带中所有帧,得到各个子带内有效音频存在的区域。

[0045] (4) 将得到的有效音频存在的区域求并集,即可得到增强后的音频信号的有效音频存在的区域。

[0046] 有益效果:本算法通过采用双阈值判别方式,设定了绝对阈值 $\eta_2 E_{noise}(i, j)$ 与保护阈值 $\eta_1 E_{noise}(i, j)$ 两个判定门限,能够避免使用单一阈值判定时出现误判的可能性,该判别方式检测的准确度高于单一阈值检测算法,能够更好地区分语音帧和噪声帧。且该检测单元针对每个子带进行检测,该检测法能够忽略噪声本身能量在不同频率段分布的差异性,使用的噪声范围更广。

[0047] 在一个实施例中,所述根据检测子单元222的检测结果,对当前帧的噪声能量估计值进行更新,并将更新后的噪声能量估计值应用于所述检测单元中,进而对有效音频存在的区域进行更新,当满足 $|E_{noise}^k(i, j) - E_{noise}^{k-1}(i, j)| \leq \sigma$ 时,则停止对第 i 帧的噪声能量估计值进行更新, σ 是预设的常数, k 为迭代次数,其中,对当前帧的噪声能量估计值进行更新的更新公式为:

[0048]

$$E_{noise}(i, j) = \begin{cases} E_{noise}(i-1, j) & V_{cross}(i, j) = 1 \\ R(i, i-1)E_{noise}(i-1, j) + R(i, i-2)E_{noise}(i-2, j) & \text{else} \end{cases}$$

[0049] 式中, $R(i, i-1)$ 表示第 i 帧与第 $i-1$ 帧的噪声能量估计值的相关系数; $R(i, i-2)$ 表示第 i 帧与第 $i-2$ 帧的噪声能量估计值的相关系数。

[0050] 有益效果:通过实时对当前帧的噪声能量估计值进行更新,进而将更新后的噪声能量估计值应用于检测子单元222之中,进一步修正有效音频区域的范围,该算法提高了在不同场合下的适应能力,即提升了鲁棒性。且在对当前帧的噪声能量估计值进行更新时,当判定结果显示第 i 帧是噪声帧时,考虑了第 $i-1$ 帧和第 $i-2$ 帧的噪声能量估计值,以及与第 i 帧的相关程度,能够更加准确地对噪声能量估计值进行更新。

[0051] 最后应当说明的是,以上实施例仅用以说明本发明的技术方案,而非对本发明保护范围的限制,尽管参照较佳实施例对本发明作了详细地说明,本领域的普通技术人员应当分析,可以对本发明的技术方案进行修改或者等同替换,而不脱离本发明技术方案的实质和范围。

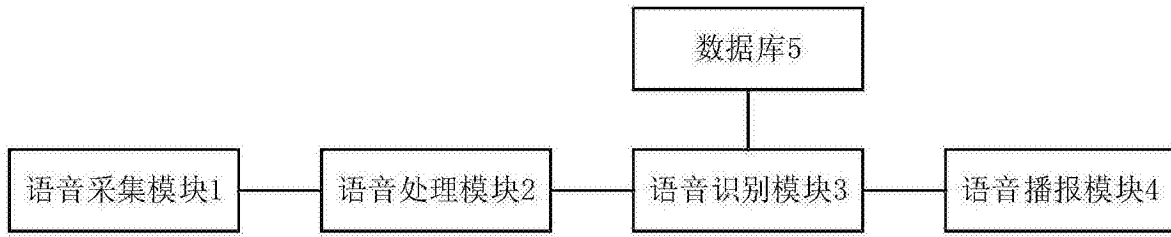


图1

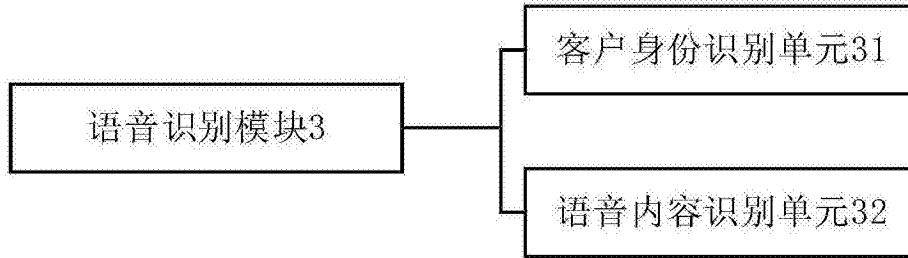


图2

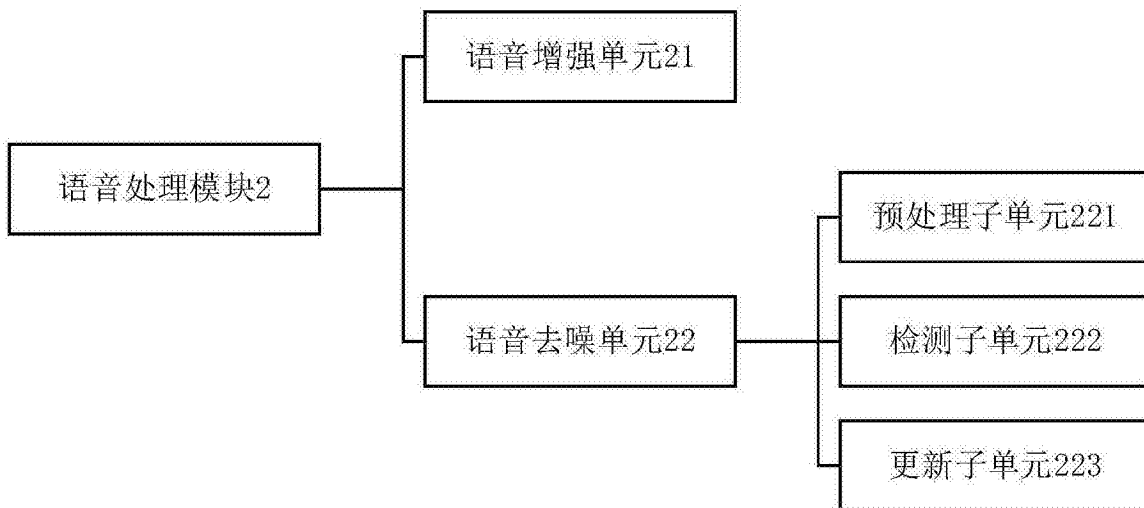


图3