

量子语音多带激励算法*

梁彦霞^{1)†} 聂敏¹⁾ 刘欣²⁾ 张美玲¹⁾ 姜静¹⁾

1)(西安邮电大学通信与信息工程学院, 西安 710121)

2)(中兴通讯西研所, 西安 710065)

(2014年1月16日收到; 2014年3月4日收到修改稿)

将经典语音多带激励(MBE)算法应用于量子领域, 提出经典信息与量子信息的对应关系, 并给出相应的信息测度方法. 对量子语音MBE的编解码算法用C语言进行仿真实验, 结果证明, 语音分解与合成前后的波形相近, 用PESQ软件客观测试语音, MOS分数为3.337.

关键词: 量子, 语音, 对应关系, 量子测度

PACS: 03.67.Hk, 03.67.Lx, 03.67.-a, 03.67.Pp

DOI: 10.7498/aps.63.120301

1 引言

量子通信作为物理学科和信息学科的交叉学科, 近年来得到了较快的发展. 量子通信领域在量子纠缠^[1-3]、量子隐形传态^[4,5]、量子中继^[6,7]、量子信令^[8-10]、量子秘密共享^[11-13]、量子压缩编码^[14]等方面累积了大量的研究基础.

在经典通信领域, 传输的信息具有一定的冗余性, 去除信息冗余以减少传输信息比特的方法称为压缩, 信息的压缩可以起到节约资源的目的. 在量子通信领域, 也要进行量子压缩, 即将表达量子信息的多个量子比特采用较少的量子比特来传输, 以达到节约传输量子比特的目的. 在过去的二十几年里, 很多学者对量子压缩编码进行了研究.

1995年, Schumacher^[15]证明了量子信息论类似于经典信息论中无噪编码的理论, 但是未进行评估. Cleve和DiVincenzo^[16]将Schumacher的编码方法进行了仿真和评估. 文献^[17]证明了在量子领域不能像经典领域那样可以完全利用当地的冗余来进行分布式压缩, 但是在某些场合, 也可以采用一定的策略来利用该冗余. 而2002年, Hayashi^[18]

提出了固定长度及变长编码算法, 找到了量子固定长度纯态源编码错误概率的最优指数, 适用于盲编码与可视编码. 2007年, 文献^[19]给出了利用 $O(n(\log^4 n))$ 步基本量子操作进行量子压缩与解压缩, 在信息保真度和块的大小上均为近似最优的, 并且文中还给出了实现该类压缩的量子门阵列模型. 2010年, 在IEEE主办的信息科学与管理工程会议上, 有学者提出了基于量子神经网络的图像压缩方法^[20]. 2014年, Feng和Zhou^[21]提出了基于量子神经模型和人工神经网络的量子反向传播(quantum reverse propagation, QBP)算法, 并用于图像压缩, 用更少的学习迭代过程得到了比BP(back propagation)算法更好的重建图像质量. 因此, 到目前为止, 已经有很多人研究了量子压缩问题, 但是并未有人将量子与语音压缩进行结合的实例.

Griffin和Lim^[22]于1988年针对传统二元模型中的缺点提出一种新的语音模型, 称为多带激励(multi-band excitation, MBE)模型. 简单的二元清/浊音模型中, 将一整段语音完全判决为清音或者浊音, 但在实际语音中, 这一段语音可能既含

* 国家自然科学基金(批准号: 61172071, 61102047)、国家科技重大专项(批准号: 2012ZX03001025-004)和陕西省教育厅科研计划专项(批准号: 11JK1016)资助的课题.

† 通讯作者. E-mail: liangyanxia201@xupt.edu.cn

有清音又含有浊音, 如果根据某个标准只判决为一种, 那么可能会在合成语音中产生嗡嗡声 [22]. MBE 模型把这一段语音划分成更小的子带, 在每个子带中进行判决, 大大接近实际语音情况, 提高了合成语音的质量. 为了进一步提高合成语音质量, 将其中分段语音的重要参数——线谱频率对 (line spectrum pair, LSP) 进行滑动平均多级分裂矢量量化, 搜索复杂度和存储空间均降低了约三分之一, 语音质量达到通信质量要求 [23].

为了将量子压缩算法在语音域进行应用, 首先要解决经典领域语音编码算法与量子通信的关系, 并建立起一套经典量子对应关系. 本文第二部分介绍量子编码算法与经典 MBE 语音编码算法; 第三部分分析二者之间的对应关系, 并给出实际对应方法及接收端的信息测度方法; 最后一部分对仿真结果进行分析并总结.

2 量子编码算法与 MBE 算法

2.1 量子编码算法

量子压缩的基本思想是用一定数量的量子比特来准确表示量子信息, 并在信道上采用数量更少的量子比特也就是压缩了的信息来传输. 用保真度 (fidelity) 描述量子信息传输的质量, 这就引入了由所有可能的量子码字所定义的典型空间的概念. 量子压缩编码算法的具体过程如下.

在量子信息通道, 用纯态序列 $\{|a\rangle, |b\rangle\}$ 来表示需要传输的信息

$$\begin{cases} |a\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ |b\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \end{cases} \quad (1)$$

假设经典信息源 X 服从一致分布, 则两个两字符号的发生概率相等 $p_a = p_b = 1/2$. 因此, 符号密度算符则为

$$\rho = p_a |a\rangle\langle a| + p_b |b\rangle\langle b| = \frac{1}{4} \begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}, \quad (2)$$

则 ρ 的特征值分别为

$$\lambda_a = (1 + 1/\sqrt{2})/2 \equiv \cos^2(\pi/8)$$

和

$$\lambda_b = (1 - 1/\sqrt{2})/2 \equiv \sin^2(\pi/8).$$

相对应的特征向量

$$|\lambda_a\rangle = \begin{pmatrix} \cos(\pi/8) \\ \sin(\pi/8) \end{pmatrix}, \quad (3)$$

$$|\lambda_b\rangle = \begin{pmatrix} \sin(\pi/8) \\ -\cos(\pi/8) \end{pmatrix}. \quad (4)$$

基于特征态 $\{|\lambda_a\rangle, |\lambda_b\rangle\}$, 密度算符 ρ 的对角形式为

$$\begin{aligned} \rho &= \begin{pmatrix} \lambda_a & 0 \\ 0 & \lambda_b \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\pi/8) & 0 \\ 0 & \sin^2(\pi/8) \end{pmatrix}. \end{aligned} \quad (5)$$

每量子比特的熵为

$$\begin{aligned} S(\rho) &= -\lambda_a \log \lambda_a - \lambda_b \log \lambda_b \\ &= -\lambda_a \log \lambda_a - (1 - \lambda_a) \log(1 - \lambda_a) \\ &\equiv 0.6008. \end{aligned} \quad (6)$$

特征态 $|\lambda_a\rangle, |\lambda_b\rangle$ 和量子符号 $|a\rangle$ 和 $|b\rangle$ 之间的交叠按下式计算:

$$\begin{aligned} |\langle \lambda_a | a \rangle|^2 &= |\langle \lambda_b | b \rangle|^2 \\ &= \cos^2(\pi/8) = \lambda_a \approx 0.8535, \end{aligned} \quad (7)$$

$$\begin{aligned} |\langle \lambda_b | a \rangle|^2 &= |\langle \lambda_b | b \rangle|^2 \\ &= \sin^2(\pi/8) = \lambda_b \approx 0.1465. \end{aligned} \quad (8)$$

因此, 以 $|\lambda_a\rangle$ 为特征态的子空间和 $|a\rangle$ 的交叠态, 比以 $|\lambda_b\rangle$ 为特征态的子空间和 $|b\rangle$ 的交叠态更大, 即 85% 对 15%.

保真度的概念定义为

$$F = \langle \psi | \rho | \psi \rangle = p_a |\langle \psi | a \rangle|^2 + p_b |\langle \psi | b \rangle|^2, \quad (9)$$

式中, $|\psi\rangle$ 为任意测量态. 这里, 因为 $p_a = p_b = 1/2$, 保真度降至

$$F = (|\langle \psi | a \rangle|^2 + |\langle \psi | b \rangle|^2)/2.$$

由 (9) 式, 显然保真度是对于测量的选择 $|\psi\rangle = |\lambda_a\rangle$ 的一个最大值, 无论 $|a\rangle, |b\rangle$ 是否是由发起人发出的, 可得到 $F = 0.853$. 因此, $\{|\psi\rangle\} = \{|\lambda_a\rangle\}$ 对应一个一维“类子空间 (likely subspace)”, 在此空间中, 任何消息符号均最强交叠.

2.2 MBE 语音编码算法

在 MBE 声码器中, 包含有基音频率 ω_0 , 清/浊音判决参数 U_k , 以及谱包络抽样 $\{A_m\}$. 其中谱包络抽样以基音频率为间隔沿频率轴均匀分布, 对于每一个谐波带来说只有一个谱幅度抽样, 这些抽样表示的是语音的平滑谱(共振峰结构). 而每一个划分的子带中的谐波个数是不一样的, 因而每个子带中的谱包络抽样也不一样. 这样, $\{A_m\}$ 的维数是变化的, 将其转换为量子比特也比较困难. 因此, 我们将 $\{A_m\}$ 转化成一组固定个数的线性预测编码(linear prediction coding, LPC) 参数 $\{a_k\}$.

LPC 参数表述了语音的短时谱包络信息, 对保证语音质量和压缩码率起着直接作用. 表达 LPC 参数的最有效方式为 LSP 参数. LSP 参数具有自然顺序、很容易与语音理解中已知的频谱特点相结合以及与共振峰频率关系密切的特性. 一个 LSP 参数的误差仅仅影响全极点模型中邻近这个参数对应频率处的语音谱, 不影响其他地方.

在改进的 MBE 算法^[24] 中, 经典信息基音频率 ω_0 , 清/浊音判决参数 U_k , 增益值 G 和 LSP 参数最终完全可以采用经典信息 0, 1 比特进行表示.

3 语音参数的量子态对应关系及测量设计

经过对量子压缩算法的研究, 量子通道传输的信息采用纯态, 其具有以下几种特征:

- 1) 纯态 $|a\rangle$ 和 $|b\rangle$ 是表示信息状态的基本元素, 二者不会同时出现;
- 2) $|a\rangle$ 和 $|b\rangle$ 均用标准正交基 $|0\rangle, |1\rangle$ 来表示;
- 3) 通过对符号密度算符 ρ 的计算和推导, 可以得到其特征值及对应的特征态;
- 4) 通过其特征态与量子符号 $|a\rangle$ 和 $|b\rangle$ 的交叠性, 方能发现特定的对应关系.

基于上述几种特征, 我们选择经典领域的 0, 1 比特与量子域的纯态 $|a\rangle$ 和 $|b\rangle$ 一一对应, 并设计实验.

将文献^[24] 中改进的 MBE 编码算法中的各个参数, 包括基音频率 ω_0 , 清/浊音判决参数 U_k , 增益值 G 和 LSP 参数的二进制序列, 用纯态 $|a\rangle$ 和 $|b\rangle$ 表示. 其中 LSP 参数采用三级矢量量化方法. 在接收端对这些状态进行测度, 从而分辨出发送的信息,

进而恢复经典的 0, 1 序列, 进行语音合成, 并对原始语音与合成语音进行对比.

然而, 由于量子态本身的特性, 如果随机选择基对量子态进行投影测量, 则量子的状态将会塌缩至该基上, 只能得到该量子态在该基上的概率, 而不能区分发送方发送的具体信息是什么, 再次的测量也无济于事. 因此, 将测量结果的保真度依赖于测量基选取的正确性, 不能准确得到测量结果, 特别是当两类基测量态出现的概率相等时. 而事实上, 量子态对于测量基来说概率相等的情形是很常见的. 本实验采用半正定算子值检测法(positive operator-valued measure, POVM).

在该检测算法中, 考虑有三个元素构成 POVM,

$$\begin{cases} E_1 = \frac{\sqrt{2}}{1 + \sqrt{2}} |1\rangle\langle 1|, \\ E_2 = \frac{\sqrt{2}}{1 + \sqrt{2}} \frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2}, \\ E_3 = I - E_1 - E_2. \end{cases} \quad (10)$$

设 Bob 接收到的状态为 $|\phi\rangle$, 对其进行 POVM $\{E_1, E_2, E_3\}$ 测量. 采用测量元素 E_1 , 可以得到如下关系:

$$\langle a|E_1|a\rangle = \left\langle 0 \left| \frac{\sqrt{2}}{1 + \sqrt{2}} |1\rangle \right. \right\rangle \langle 1|0\rangle = 0, \quad (11)$$

$$\begin{aligned} \langle b|E_1|b\rangle &= \left\langle 0 \left| \frac{\sqrt{2}}{1 + \sqrt{2}} \frac{(|0\rangle - |1\rangle)(\langle 0| - \langle 1|)}{2} \right. \right\rangle \left. \right\rangle \\ &= \frac{\sqrt{2}}{2(1 + \sqrt{2})} \neq 0, \end{aligned} \quad (12)$$

可见, 此时得到结果 E_1 的概率为 0. 如此, Bob 可知, 他收到的比特为 $|b\rangle$. 也就是说, 如果测量的结果 E_1 , 则 Bob 收到的是 $|b\rangle$; 同样, 若测量结果 E_2 , 则 Bob 可知收到的比特为 $|a\rangle$.

由此实现了经典信息与量子信息的对应, 以及接收端检测信号的方式, 从而建立了一套通过量子信道传输语音信息的量子 MBE 语音编码算法.

4 结果与分析

在 MBE 语音编解码系统中, 采用 8 kHz 采样, 帧长 20 ms, 每帧 160 个抽样. 在此前提下, 在一帧中基音频率 ω_0 , 清/浊音判决参数 U_k , 增益值 G 和 LSP 参数在经典通信域分别占用 10, 6, 5 和 27 bits

(三级量化, 每级需要 9 bits), 因此传输一帧语音需要 48 bits, 并对应为 48 个量子态. 若经典信息为 0, 则认为量子信道需传输纯态 $|a\rangle$; 若为 1, 则传输纯态 $|b\rangle$. 本系统不考虑具体传输的方式, 只考虑发送端、接收端的量子态变换, 并用 POVM 进行测量, 以观察是否可以实现语音的恢复, 并考核语音质量. 原始语音与合成语音的图谱见图 1.

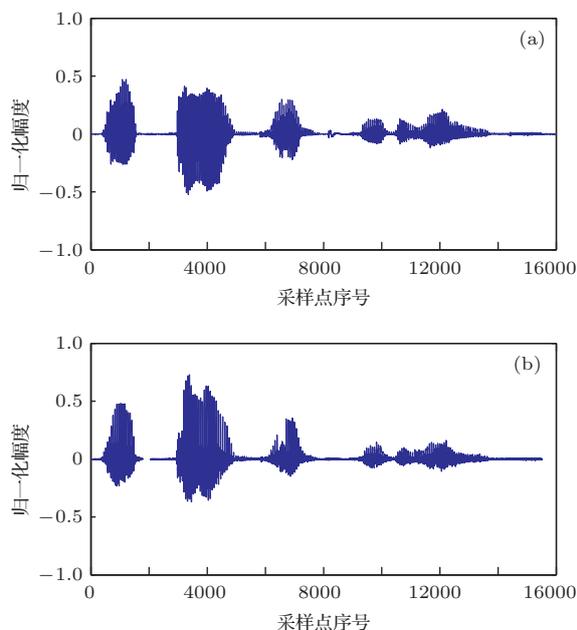


图1 原始语音与合成语音图谱比较 (a) 原始语音; (b) 合成语音

从图 1 中可以看出, 合成后的语音与原始语音的波形接近, 清/浊音判决大致相近. 试听结果显示, 可以听清楚语音内容, 并无杂音或其他信号. 将结果用客观语音质量评估 (perceptual evaluation of speech quality, PESQ) 的 VC 程序进行测试, 得到的 MOS 分数为 3.337, 基本达到了通信的质量.

5 结 论

本文提出了经典与量子信息的对应关系, 并且给出了接收端对量子信息的测度方法. 该方法能够正确恢复出语音, 且语音清晰易懂. 在客观语音质量测试软件的测试下, 达到了通信质量. 本文将经典语音低速率 MBE 编码算法应用于量子领域, 是在特定业务情形下量子与经典通信的一种融合

研究.

参考文献

- [1] Zhou N R, Zeng B Y, Wang L J, Gong L H 2010 *Acta Phys. Sin.* **59** 2193 (in Chinese) [周南润, 曾宾阳, 王立军, 龚黎华 2010 物理学报 **59** 2193]
- [2] Ji Y H, Hu J J, Hu Y 2012 *Chin. Phys. B* **21** 110304
- [3] Zhou N R, Zeng G H, Gong L H, Liu S Q 2007 *Acta Phys. Sin.* **56** 5066 (in Chinese) [周南润, 曾贵华, 龚黎华, 刘三秋 2007 物理学报 **56** 5066]
- [4] Zhou X Q, Wu Y W 2012 *Acta Phys. Sin.* **61** 170303 (in Chinese) [周小清, 鄂云文 2012 物理学报 **61** 170303]
- [5] Wang M Y, Yan F L 2011 *Chin. Phys. B* **20** 120309
- [6] Djordjevic I B 2011 *IEEE Photon. J.* **2** 81
- [7] Pei C X, Yan Y, Liu D, Han B B, Zhao N 2008 *Acta Photon. Sin.* **37** 2422 (in Chinese) [裴昌幸, 阎毅, 刘丹, 韩宝彬, 赵楠 2008 光子学报 **37** 2422]
- [8] Nie M, Zhang L, Liu X H 2013 *Acta Phys. Sin.* **62** 230303 (in Chinese) [聂敏, 张琳, 刘晓慧 2013 物理学报 **62** 230303]
- [9] Zhu W, Nie M 2013 *Acta Phys. Sin.* **62** 130304 (in Chinese) [朱伟, 聂敏 2013 物理学报 **62** 130304]
- [10] Xue L, Nie M, Liu X H 2013 *Acta Phys. Sin.* **62** 170305 (in Chinese) [薛乐, 聂敏, 刘晓慧 2013 物理学报 **62** 170305]
- [11] Zhu Z C, Zhang Y Q, Fu A M 2012 *Chin. Phys. B* **21** 010307
- [12] Du J Z, Sun Y, Qin S J, Wen Q Y, Zhu F C 2008 *Acta Phys. Sin.* **57** 4694 (in Chinese) [杜建忠, 孙莹, 秦素娟, 温巧燕, 朱甫臣 2008 物理学报 **57** 4694]
- [13] Xiao H L, Guo W M, Wang X 2012 *Chin. Phys. Lett.* **29** 110301
- [14] Liang Y X, Nie M 2013 *Acta Phys. Sin.* **62** 200305 (in Chinese) [梁彦霞, 聂敏 2013 物理学报 **62** 200305]
- [15] Schumacher B 1995 *Phys. Rev. A* **51** 2738
- [16] Cleve R, DiVincenzo D P 1996 *Phys. Rev. A* **54** 2636
- [17] Ahn C, Doherty A C 2006 *IEEE Trans. Inform. Theory* **52** 4349
- [18] Hayashi M 2002 *Phys. Rev. A* **66** 032321
- [19] Reif J H, Chakraborty S 2007 *Inform. Comput.* **205** 967
- [20] Li H F, Li M 2010 *Conference of Information Science and Management Engineering (ISME)* Xi'an, China, August 7-8, 2010 p567
- [21] Feng Q G, Zhou H Y 2014 *J. Electr. Engineer.* **12** 197
- [22] Griffin D W, Lim J S 1988 *IEEE Trans. Speech Sign. Process.* **36** 1223
- [23] Liang Y X, Yang J W, Li Y, Nie M, Liu X 2012 *J. Jilin Univ. (Engineering and Technology Edition)* **42** 223 (in Chinese) [梁彦霞, 杨家炜, 李焯, 聂敏, 刘欣 2012 吉林大学学报 (工学版) **42** 223]
- [24] Liang Y X, Yang J W, Li Y 2010 *J. Xidian Univ.* **37** 417 (in Chinese) [梁彦霞, 杨家炜, 李焯 2010 西安电子科技大学学报 **37** 417]

Quantum speech multi-band excitation algorithm*

Liang Yan-Xia^{1)†} Nie Min¹⁾ Liu Xin²⁾ Zhang Mei-Ling¹⁾ Jiang Jing¹⁾

1) (*School of Telecommunication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China*)

2) (*Xi'an Research and Development Center, ZTE Corporation, Xi'an 710065, China*)

(Received 16 January 2014; revised manuscript received 4 March 2014)

Abstract

Multi-band excitation (MBE) algorithm in classical communication is applied to quantum communication. The corresponding relation between classical information and quantum information is proposed, and the quantum measurement of information is presented. Simulations of quantum MBE algorithm by C language programming demonstrate that the wave of synthesized speech is similar to the original one. An MOS score of 3.337 is obtained from the Perceptual Evaluation of Speech Quality software.

Keywords: quantum, speech, correspondence, quantum measurement

PACS: 03.67.Hk, 03.67.Lx, 03.67.-a, 03.67.Pp

DOI: [10.7498/aps.63.120301](https://doi.org/10.7498/aps.63.120301)

* Project supported by the National Natural Science Foundation of China (Grant Nos. 61172071, 61102047), the National Science and Technology Major Project of the Ministry of Science and Technology of China (Grant No. 2012ZX03001025-004), and the Special Scientific Research Foundation of the Education Department of Shaanxi Province, China (Grant No. 11JK1016).

† Corresponding author. E-mail: liangyanxia201@xupt.edu.cn