Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 7 (22): 3567-3571, 2007
ISSN 1812-5654

Adaptive Quantum Lossless Compression

Essam Al-Daoud
Department of Computer Science, Faculty of Science and Information Technology,
Zarka Private University, Jordan

Abstract: A new adaptive quantum lossless compression algorithm is introduced, the suggested algorithm does not need a priori estimation of probabilities and it is more useful in real applications. The main idea of the proposed algorithm is that the corresponding probabilities of the symbols are assumed to be identical and after each iteration; all the corresponding probabilities are updated. These probabilities will become closer to the actual distribution after few iterations. Moreover the complexity of the proposed algorithm can be reduced if we consider the general properties of the given data

Key words: Entropy, quantum computer, quantum compression, density matrix, adaptive algorithm

INTRODUCTION

Quantum computer is any device for computation that makes direct use of distinctively quantum mechanical phenomena. In a quantum computer, the data is measured by qubits (quantum bits). There has been much recent interest in the subject of quantum information processing. Quantum information is a natural generalization of classical information. It is based on quantum mechanics, a well-tested scientific theory in real experiments. Although quantum computation and communication are still in its infancy, experiments have been carried out in which quantum computational operations were executed on a very small number of qubits. Research in both theoretical and practical areas continues at a frantic pace and many national government and military funding agencies support quantum computing research to develop quantum computers for both civilian and national security purposes (Panthong *et al*., 2005). D-Wave Systems demonstrated on February 13 and 15th 2007 what is claimed to be the worlds first commercial quantum computer by using 16 qubits (www.dwavesys.com).

Quantum information differs from classical information in several respects such as the states teleportation, the states cannot generally be read or duplicated without disturbance (no cloning theorem), one state can exist in superposition of all possible states at once and there are statistical correlations predicted by quantum physics for measurements on two entangled particle systems. The ability to manipulate quantum information enables us to perform tasks that would be unachievable in a classical context, such as unconditionally secure transmission of information, quantum authentication, quantum digital signature and solving the hard problems in polynomial time (Al-Daoud, 2007).

It may be very advantageous to decrease, where possible by compression methods, the number of qubits used for quantum communication and storage. This study introduces a new compression method without using the statistical distribution of the given sequence of the qubits and analogy to the adaptive Huffman code.

CLASSICAL LOSSLESS COMPRESSION

Information theory is generally considered to have been founded in 1948 by Claude Shannon in his seminal work, A Mathematical Theory of Communication. He established the two core results of classical information theory in his landmark. The two central problems that he solved were: the amount of the compression done on a message and the necessary rate communicated reliably over a noisy channel. Both problems concern redundancy. Shannon introduced a new entropy definition in the theory of information, it is of the form (Shannon, 1948):

$$H=-\sum p_i \log p_i$$

where:
p_i = The probability.

Applications of fundamental topics of information theory include lossless data compression (e.g., ZIP files), lossy data compression (e.g., MP3s, MPEG and JPEG) and channel coding (e.g., for DSL lines).

The above entropy definition can be used to determine the theoretical lossless compression lower bound or the compression rate. The compression rate is the ratio between the length of an uncompressed string

and the length of the compressed (binary) string. There are two types of the universal data compression in classical domain: the first type is two pass compression algorithms such as Run-length Encoding, Huffman coding and Arithmetic coding. The second type is one pass compression algorithms such as Adaptive Huffman coding, LZ77 and LZ78. Ziv and Lempel present a simple linear time lossless compression algorithm having an asymptotic compression rate approaching the sources entropy; that is allowing a string of length n to be losslessly compressed to a bit string of length asymptotic approaching H(p) n for large n. In the first pass, they use a parsing scheme to encode the source string into unique prefixes.

QUANTUM LOSSLESS COMPRESSION

Quantum lossless compression is one of the important directions of the quantum information processing which starts from the thermodynamic entropy. Gibbs defined The thermodynamic entropy S after earlier work by Boltzmann as follows:

$$S=-K_B\sum p_i \ln p_i$$

Assume that the underlying ensemble is $\Sigma = \{P, X\}$, where, X is the set of all symbols $X = \{|\Psi_1>, |\Psi_2>,\ldots|\Psi_n>\}$ and P is the set of corresponding probabilities, hence the Gibbs entropy translates over almost unchanged into the world of quantum physics to give the von Neumann entropy formula (Jozsa *et al*., 1998):

$$S = Tr(\rho \ln \rho)$$

where:

ρ = The density matrix of the quantum mechanical system defined as follows:

$$\rho=\sum_{i=1}^{n} pi|\psi_i><\psi_i|$$

Benjamin Schumacher is a US theoretical physicist, working mostly in the field of quantum information theory. He discovered a way of interpreting quantum states as information. He came up with a way of compressing the information in a state and storing the information in a smaller number of states. This is now known as Schumacher compression. This was the quantum analog of Shannon's noiseless coding theorem and it helped to start the field known as quantum information theory.

Schumacher quantum lossless compression algorithm can be described as follows:

- Select a typical sub message $|\Psi_{typ}>$ and ignore the rest of the message, where the typical sub messages is in the subspace that spanned by $\{|\Psi_1>, |\Psi_2>,\ldots|\Psi v>\}$ and $v = 2^{nS}$.
- Apply a unitary change of basis U that takes $|\Psi_{typ}>$ to a state of the form:

$$U|\Psi_{typ}> = |\Psi_{comp}>|0>\otimes|0>\otimes\ldots\otimes|0>$$

- Send $|\Psi_{comp}$

Schumacher decompression can be done by appending the zeros to $|\Psi_{comp}>$ and applying U^{-1}. Moreover Schumacher proves that the compressed qubits equals to $m(S+\delta)$,where, m is the length of typical sub message (Schumacher, 1995).

Braunstein and others introduce a quantum analog of Huffman coding by using divide and conquer. Firstly, they divide the messages into pairs and apply a merging procedure to each pair. The merging effectively reduces the total number of messages from 2^r to 2^{r-1}. This process can be repeated. Therefore, after r applications of the merging procedure, we obtain a single tape containing all the messages (in addition to the various length tapes containing the length information). Let us introduce a message tape, for simplicity, we simply denote $|0\cdots 0h_i>$ by $|h_1>$, etc.

$$\begin{aligned} &|h_1>|l_1>|h_2>|l_2>|0>_{tape}\\ \text{swap}\rightarrow\ &|0>|l_1>|h_2>|l_2>\ |0\cdots 0h_1>_{tape}\\ \text{shift}\rightarrow\ &|0>|l_1>|h_2>|l_2>\ |\ h_1 0\cdots 0>_{tape}\\ \text{swap}\rightarrow\ &|0>|l_1>|0>|l_2>\ |\ h_1 0\cdots h_2>_{tape}\\ \text{shift}\rightarrow\ &|0>|l_1>|0>|l_2>\ |\ h_1\ h_2 0\cdots 0>_{tape}\end{aligned}$$

In general, at the end the encoder obtains:

$$|l_1>|\ l_2>\cdots|\ l_n>|\ h_1\ h_2\cdots h_n 0\cdots 0>_{tape}$$

The encoder truncates the message tape: he keeps the first $N(L+\delta)$ qubit in the message tape (Braunstein *et al*., 2002).

Bostroem and Felbinger (2002) develop a general framework for variable-length quantum messages in close analogy to the classical case. They show that the lossless compression of an ensemble of messages is bounded from below by its von-Neumann entropy and it is possible to reduce the number of qubits passing through a quantum channel even below the von Neumann entropy by adding a classical side channel. they give an explicit communication protocol that realizes lossless and instantaneous quantum data compression.

Bennett gave a constructive method for doing Schumacher compression. He observed that the

Schumacher compression can be done by a unitary mapping to a basis for which the density matrix ñ is diagonal followed by certain combinatorial. We can perform the combinatorial by ordering the basis states first by the number of ones (from smallest to largest) that are in the binary expansion of the bits and then refining this order by a lexical sort of the binary expansion of the bits (Reif and Chakraborty, 2007).

THE PROPOSED ALGORITHM

Assume that Alice likes to send a stream of compressed symbols (characters) to Bob through a quantum channel, the symbols before the compressing can be written as $M = |x_1\rangle|x_2\rangle\ldots|x_m\rangle$ where, $|x_i\rangle \epsilon X, \forall\, i = 1, 2..., m$ and m is the message length. Alice does not know what is the coming symbol, thus she does not know the probability distribution of the message. The following steps can be used to compress a stream of data [some notations and steps are borrowed from (Bostroem and Felbinge, 2002)]:

1. Alice assumes that all the symbols (characters) have identical probability, i.e., Alice has the ensemble $\Sigma = \{P, X\}$, where, X is the set of all symbols $X = \{|\Psi_1\rangle, |\Psi_2\rangle,\ldots|\Psi_n\}$, P is the set of corresponding probabilities which is initially equals to {1/n, 1/n,...1/n} and n is the number of symbols. Moreover each symbol can be represented by r qubit.

2. $j = 1$, $Counter_i = 0, \forall\, I=1, 2..., n.$

3. Alice prepares the subset $L = \{|\Psi_1\rangle,|\Psi_2\rangle,\ldots|\Psi_d\} \subseteq X$ of linear independent vectors, where these vectors are selected and ordered by the highest probability.

4. Alice finds the orthornormal vectors $B=\{|w_1\rangle,|w_2\rangle,\ldots,|w_d\rangle\}$ by performing a Gram- Schmidt orthornormalization on the list L. thus the set B can be defined as follows:

$$|w_1\rangle = |x_1\rangle,$$

$$|w1\rangle = C_i\left[\sum_{k=1}^{i-1} |\, w_k \rangle\langle w_k \,|\right] |\, x_i \rangle,\ i = 1, 1,..., d$$

5. Alice calculates the unitary matrix as follows:

$$C=\sum_{i=1}^{d} |\, Z(i-1) \rangle\langle w_i \,|$$

where, the state $|Z(i-1)\rangle$ must be represented in a neutral-prefix, which means that the number of qubit required to represent the state $|Z(i-1)\rangle$ is equal to the number of qubit required to represent the longest symbol(character). Zeros add to the left, for example if the longest symbol needs 5 qubit then $|Z(3)\rangle=|00011\rangle$.

6. Alice picks up the next symbol and encodes it as follows:

$$|c(y)\rangle = C|y\rangle, \text{ where, y is the picked symbol}$$

7. Alice calculates the base length L of $|c(y)\rangle$, where, L is the longest component of the state $|c(y)\rangle$ as defined by (Bostroem and Felbinger, 2002).

8. Alice truncates the message to L qubits by removing r-L leading qubits.

9. Alice sends the truncated message through a quantum channel and sends its base length L through a classical channel (adaptive huffman can be used to encode the bases length).

10. j = j+1, if j>m then stop.

11. $Counter_t = Counter_t + 1$, where, t indicates the t^{th} symbol in the set X such that $|\Psi_1\rangle=|y\rangle$.

12. If $Counter_t = 1$ then go to step 6

13. If $Counter_t > 1$ then update the set of corresponding probabilities as follows:

$$\begin{cases} P(|\psi_t\rangle) = P(|\psi_t\rangle) + 1/n \\ P(|\psi_k\rangle) = P(|\psi_k\rangle) - (n-1)/n \quad \forall k \neq t \end{cases}$$

14. Go to step 3.

Bob receives the stream of base length from the classical channel and the stream of compressed symbols from the quantum channel and decompresses the message by using the following steps:

1-5. Bob performs the steps 1-5 as Alice does.

6 Bob decompresses the base length L of the next symbol.

7. Bob adds (r-L) zeros to the right of the received state from the quantum channel, call it $|\gamma\rangle$

8. Bob decodes (decompress) the state $|\gamma\rangle$ as follows:

$$|y\rangle = |d(\gamma)\rangle = D|\gamma\rangle = C^{-1}|\gamma\rangle = C^{t}|\gamma\rangle$$

9. j = j+1, if j>m then stop.

10. $Counter_t = Counter_t + 1$, where, t is indicate to the t^{th} symbol in the set X such that $|\Psi_t\rangle=|y\rangle$

11. If $Counter_t = 1$ then go to step 6

12. If $Counter_t > 1$ then update the set of corresponding probabilities as follows:

$$\begin{cases} P(|\psi_t\rangle) = P(|\psi_t\rangle) + 1/n \\ P(|\psi_k\rangle) = P(|\psi_k\rangle) - (n-1)/n \quad \forall k \neq t \end{cases}$$

13. Go to step 3.

We can reduce the complexity of the above algorithm if we consider that the corresponding probabilities will become closer to the actual distribution after few iterations. Thus we can ignore the recalculation of the unitary matrix and jump directly to the step 6, or the calculation can be postponed until critical changing is occurred in the set of the corresponding probabilities.

THE SIMULATION OF THE PROPOSED ALGORITHM

Let us first discuss an explicit example to demonstrate the steps of the suggested algorithm. Assume that Alices source message set is:

$$X=\{|a\rangle,|b\rangle, |c\rangle, |d\rangle,|e\rangle,|f\rangle, |g\rangle, |h\rangle\}$$

whose elements are in the mixed states and given by:

$$a=1/\sqrt{15}\,|00\rangle+1/\sqrt{15}\,|10\rangle+2/\sqrt{15}\,|01\rangle+1/\sqrt{15}\,|11\rangle=1/\sqrt{15}(1121)^t$$

$$b=1/\sqrt{15}\,|10\rangle+1/\sqrt{15}\,|01\rangle+2/\sqrt{15}\,|11\rangle+1/\sqrt{15}(01121)^t$$

$$c=1/\sqrt{7}\,|00\rangle+2/\sqrt{7}\,|10\rangle+1/\sqrt{7}\,|11\rangle=1/\sqrt{7}(1201)^t$$

$$d=1/\sqrt{25}\,|00\rangle+3/\sqrt{25}\,|10\rangle+1/\sqrt{25}\,|01\rangle+3/\sqrt{25}\,|11\rangle=1/\sqrt{25}(1313)^t$$

$$e=1/\sqrt{30}\,|00\rangle+2/\sqrt{30}\,|10\rangle+3/\sqrt{30}\,|01\rangle+3/\sqrt{30}\,|11\rangle=1/\sqrt{30}(1\,2\,3\,3)^t$$

$$f=1/\sqrt{15}\,|10\rangle+2/\sqrt{15}\,|01\rangle+3/\sqrt{15}\,|11\rangle+1/\sqrt{15}(0\,1\,2\,3)^t$$

$$g=4/\sqrt{18}\,|00\rangle+2/\sqrt{18}\,|10\rangle+1/\sqrt{18}\,|01\rangle+1/\sqrt{18}\,|11\rangle=1/\sqrt{18}(4211)^t$$

$$h=1/\sqrt{6}\,|10\rangle+2/\sqrt{6}\,|01\rangle+1/\sqrt{6}\,|11\rangle+1/\sqrt{6}(0\,1\,2\,1)^t$$

Now consider that Alice's message is M =| eebbd> =|e>|e>|b>|b>|d> and she likes to send it through a quantum channel. Thus Alices linear independent vectors are:

$$L = \{a, b, c, f\}\subseteq X$$

and the orthornormal vectors B = $\{|w_1\rangle,|w_2\rangle,|w_3\rangle,|w_4\rangle\}$ where:

$$w_1 = (0.3780\ 0.3780\ 0.7559\ 0.3780)^t$$
$$w_2 = (-0.4583\ 0.1833\ -0.2750\ 0.8250)^t$$
$$w_3 = (0.4291\ 0.7261\ -0.5281\ -0.0990)^t$$
$$w_4 = (0.6804\ -0.5443\ -0.2722\ 0.4082)^t$$

Hence the first unitary matrix is:

$$C_1 = \begin{bmatrix} 0.3780 & 0.3780 & 0.7559 & 0.3780 \\ -0.4583 & 0.1833 & -0.2750 & 0.8250 \\ 0.4291 & 0.7261 & -0.5281 & -0.0990 \\ 0.6804 & -0.5443 & -0.2722 & 0.4082 \end{bmatrix}$$

Now Alice can compress the first character (which consists from two qubits) as follows:

$$C1^*|e\rangle = C1^*\ 1/2/\sqrt{15}\ (1\,2\,3\,3)^t = (0.8281\ 0.2845\ 0\ 0)$$
$$= 0.8281|00\rangle + 0.2845|10\rangle$$

Alice truncates the state to the base length L = 1 qubit and obtains $0.8281|0\rangle+ 0.2845|1\rangle$. She sends the left qubit through the quantum channel. Bob can decompress the received data by calculating C_1, adding a qubit in the state $|0\rangle$ to the right of the received qubit and applying:

$$C_1^t\ (0.8281|00\rangle + 0.2845|10\rangle) = 1/\sqrt{15}(1\,2\,3\,3)^t = |e\rangle$$

To compress the second character, Alice updates the set of corresponding probabilities and finds the new linear independent vectors, where these vectors are selected and ordered by the highest probability as follows:

$$L = \{e, a, c, f\}\subseteq X$$

Alice calculates C_2 and compresses the second character as follows:

$$C_2^*|e\rangle = C_2^*\ 1/\sqrt{15}\ (1\,2\,3\,3)^t = (0.8756\ 0\ 0\ 0)$$
$$= 0.8756|00\rangle$$

Alice truncates the state to the base length L = 0 qubit. In this case there are no qubits left at all, so she sends nothing through the quantum channel and sends "0" through the classical channel (note that: the classical bits are cheaper than the quantum qubits, where, n qubits contain information equivalent to 2^n bits (Lee *et al*., 2002)). Bob receives the classical information 0. In this case he has to prepare two qubits in the state $|00\rangle$ and apply the decoder C_2^t although Alice dropped the coefficient 0.8756, Bob can find the correct character by comparing and scaling the obtained vector. Alice can compress the rest of her message by applying the same process.

Table 1 shows that where random data and real application data are used. The proposed algorithm is coded in Matlab™ 7.0 and is run in a PC with Pentium 4 microprocessor, 2.6 GH and 256 MB RAM. The

Table 1: The numerical simulation of the new algorithm

No. of qubits				
Before compression	After compression	Data source	Ratio (%)	Time (sec)
500	366	Random data	73.20	0.01
3000	2151	Random data	71.70	0.06
500	292	From a text	58.40	0.01
3000	1690	From a text	56.33	0.06
500	280	From an image	56.00	0.01
3000	1624	From an image	54.13	0.06

compression ratio is calculated with the division of compressed size by uncompressed size *100. So, lower is better.

The suggested algorithm is the first quantum lossless compression algorithm that works without priori estimation of probabilities; all the previous algorithms require statistical knowledge. Hence the previous algorithms require collecting the whole data before the compression process begins. In the real applications the suggested algorithm works better because some symbols tend to be repeated and the whole data are often not available.

CONCLUSION

The previous quantum compression lossless algorithms require the statistical knowledge which is often not available such as live audio and video. Even when the data is available, some quantum application does not allow performing the measurement more than one time. The first adaptive quantum compression algorithm without loss of information is introduced, where the source message is not known to the sender, the suggested protocol can be used for both online quantum communication and storage of quantum data.

REFERENCES

Al-Daoud, E., 2007. Unconditionally secure quantum payment system. Int. J. Applied Math. Comput. Sci., 4: 566-569.

Bostroem, K. and T. Felbinge, 2002. Lossless quantum data compression and variable length coding. Phys. Rev. Lett. A., 65: 032313.

Braunstein, S.L., C.A. Fuchs, D. Gottesmann and H.K. Lo, 2000. A quantum analog of huffman coding. IEEE Transactions Information Theory, 46: 1644-1649.

Jozsa, R., M. Horodecki, P. Horodecki and R. Horodecki, 1998. Universal quantum information compression. Physical Rev. Lett., 81: 1714-1717.

Lee, H., D. Ahn and S.W. Hwang, 2002. Dense coding in entangled states. Phys. Rev. A., 66: 024304.

Panthong, P. *et al.*, 2005. Experimental free space quantum key distribution. 4th International Conference on Optical Communication and Networks, Thailand, 14-16, ICOCN., pp: 159-161.

Reif, J.H. and S. Chakraborty, 2007. Efficient and exact quantum compression. J. Inform. Comput., 205: 967-981.

Schumacher, B., 1995. Quantum coding. Phys. Rev. Lett. A., 51: 2738-2747.

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Technol. J., 27: 623-656.