Predictive Coding for Efficient Host-Device Communication in a Pneumatic Force-Feedback Display

Christoph W. Borst
Center for Advanced Computer Studies, University of Louisiana at Lafayette
E-mail: cborst@cacs.louisiana.edu

Abstract

We investigate predictive coding for reducing the amount of data communicated between a haptic controller and a host. This allows increased update rate, which potentially improves quality even if coding is lossy. A low-order predictive coding is investigated for a pneumatic force display. Due to human and device characteristics, some compression is possible without loss, although the technique is lossy in general. Lossy uniform and nonuniform quantizers are also investigated. An experiment was conducted to determine how much data reduction is possible before compression artifacts become detectable to users.

1. Introduction

Force feedback quality depends on achievable update rates, which depend on the communication bandwidth between a haptic controller and a host. More efficient use of this bandwidth supports increased update rates, increased number of sensors and actuators, increased resolutions, and increased efficiency for data storage. This paper describes a compression system developed for an application of a pneumatic force-feedback glove. Compression of sensor and force values, combined with other protocol changes, resulted in over doubling of update rate and in a corresponding reduction in latencies, as summarized in [1, 2]. Similar techniques have been discussed in [3, 4], but no evaluation of effects on users was conducted. Although compression can be designed by minimizing a quantitative distortion metric, design and evaluation based on user perception is needed to ensure good results.

2. Encoding Method

In predictive coding, an encoder and decoder both compute an identical prediction of the next sample based on past decoded samples. The encoder encodes the difference between the prediction and the actual sample value and the result is communicated rather than the sample itself. For example, consider the reporting of samples from a single-sensor haptic device to a host machine. Let x_i be the ith sample to be reported, and let x'_i be ith decoded value. Assume $x'_0=x_0$. In a simple first-order differential encoding, the predictor is simply the last decoded sample, so the device sends $\Delta_i = e(q(x_i-x'_{i-1}))$ for each subsequent sample and the host computes $x'_i=x'_{i-1}+e^{-1}(\Delta_i)$. Here, $q()$ is a quantization function, $e()$ maps quantized value to an encoded value, and $e^{-1}()$ is its inverse. In practice, $e()$ and $e^{-1}()$ can be computed by table lookup, and $e(q())$ can be viewed as a single operation.

This method is promising when per-cycle differences are small, since differences can then be represented by fewer bits than needed for direct representation of samples. An occasional large difference can be communicated by a sequence of multiple small differences or by nonuniform quantization functions that represent large differences at low resolution and small differences at high resolution. Per-cycle differences for haptic devices tend to be small for multiple reasons. Limited human motion speed results in limited per-cycle changes in sensor readings. Low-pass filtering to reduce sensor noise limits per-cycle differences. Depending on electrical and physical behavior of actuators and on virtual couplings in force rendering algorithms, supporting large per-cycle changes may not be useful.

3. Design and Evaluation

A first-order differential encoder was designed and evaluated with the Rutgers Master II (RM) force-feedback glove [5]. Its low-pass noise filtering and low mechanical bandwidth of pneumatic actuators make it an excellent candidate for the encoding. For clarity, only index finger piston displacement is considered here, although all sensor values and force values were similarly encoded in [2]. Piston displacement sensor readings were filtered with a low-pass filter and calibrated by a piecewise cubic curve as in [2]. For this paper, communication rate was fixed at a stable 460 updates per second. Uncompressed samples were communicated at 16-bit resolution. For experiments, compressed samples were padded to also occupy 16 total bits. Haptic rendering was synchronized with the communication loop for round-trip latency of 2.2 ms. Maximum piston pressure was set to 40 psi (276 kPa).

3.1 Measured Difference Magnitudes

The first step for encoder design was to determine the range of differences required. We sampled data from three users intentionally making fast movements spanning the piston range, for 15 seconds per user. Sample values ranged from 1837 to 61316, making use of the full 16-bit range. First and second order differences were computed (see Figure 1). The largest-magnitude first-order difference was 4680 and the largest-magnitude second-order difference was 2413. Furthermore, 99% of first-order differences had magnitude below 2761 and 99% of second-order differences had magnitude below 838.

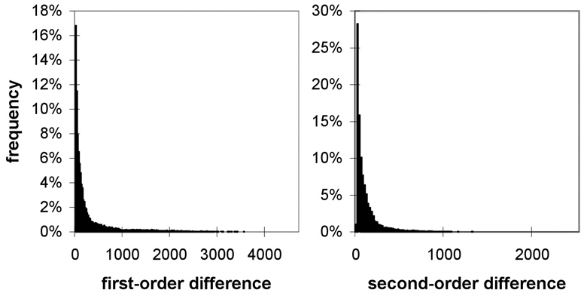

Fig. 1. Histograms for magnitudes of first-order and second-order sample differences

Although the proposed method is lossy in general, some lossless compression of these data is possible due to limited per-cycle changes. Note a 14-bit value can be used to represent the range of integers [-8192, 8191], fully capturing the range of observed first-order differences at full resolution. A first-order differential encoding can reduce the sampled data by 2 bits per sample without loss. This also extends to lower resolutions, e.g., if 8-bit resolution had been used for raw samples, lossless differential encoding would require 6 bits. Another bit reduction is possible if reduced accuracy in one percent of samples is tolerable. For a second-order differential encoding, 3 bits could be saved without loss, or 5 if reduced accuracy in one-percent of samples is tolerable.

Higher update rates can be expected to produce smaller differences. For lower rates, a maximum per-cycle difference can be estimated. For example, if update rate is halved, largest magnitude difference can be estimated to double, and one more bit is needed.

Table 1. Results of 2AFC procedure for investigating required difference range

bits required	9	10	11	12	13	14	15
total observations	3	4	13	17	12	9	2
correct responses	3	4	12	6	7	3	0

3.2 Estimation of Minimum Range to Encode

We conducted user-based evaluations of first-order differential encodings. A preliminary experiment estimated maximum encoded difference required to avoid perceptible artifacts. Our design was a 2-alternative temporal forced-choice method with tracking. Participants were three expert users experienced with force-feedback gloves, each having an understanding of haptic compression artifacts.

Each stimulus was a spring force simulated with the index finger piston. When piston displacement was below a set threshold, commanded force was proportional to the difference between the threshold and actual displacement. The "spring constant" was set high, but not so high to produce instability.

In each of 20 trials, two stimuli were presented – one rendered without compression and one using first-order differences. We refer to them as the "standard spring" and the "compressed spring". These were presented in random order for 10 seconds each. The user was required to indicate which of the two springs used compression. Experiment software indicated to users whether or not their choices were correct.

In the first trial, encoded differences were clamped to the range [-256, 255], a 9-bit range. When a user correctly identified the compressed spring, the range was doubled. For each incorrect identification, the range was halved. This was to concentrate trials in an area where detecting compression was difficult.

Table 1 shows pooled data from this experiment. Note the natural rate of correct responses when there is no detectable difference is 50%. We estimate that a 12-bit encoding would not produce detectable artifacts, based on inspection of the table and earlier data that showed over 97% of samples would be communicated without loss even for users intentionally making fast movements. Alone, this experiment is too weak for the conclusion to be strong. However, it is supported by the next experiment, which limited range to be no larger than the 12-bit range and found encodings for which users could not reasonably detect compression.

3.3 Evaluation of Quantization Functions

3.3.1 Compression Scheme. This section describes an experiment to evaluate further compression using lossy quantization in a first-order differential encoder. Range was limited to represent no differences outside of the 12-bit range [-2048, 2047] chosen in Section 3.2. So, encodings evaluated here only span at most about 3% of the total range of possible differences between two arbitrary 16-bit values.

One technique investigated for lossy quantization was to eliminate low-order bits in the differences. In other words, this was a reduction in resolution, e.g., an 8-bit uniform quantizer/encoder represented only every 16th integer in the range [-2048, 2047].

The other technique was nonuniform quantization to represent small differences at higher resolution than large differences. Specifically, for an *n*-bit scheme, we first generated a decoding table that mapped 2^n-1 table indices to the quantized values they represented using the decoding function:

$$e^{-1}(k) = \begin{cases} -s(b^{2^{n-1}-k}-1), & 0 < k < 2^{n-1} \\ s(b^{k-2^{n-1}}-1), & 2^{n-1} \le k < 2^n \end{cases}$$

where b is a user-chosen base above 1 and

$$s = \frac{2047}{b^{2^{n-1}-1}-1}$$

Quantizing/encoding mapped differences to the closest value in this table to find its index (the index being the value to transmit). This was also done with a table. We generated encodings for n ranging from 2 to 8 and b-values being (in order) 1.1, 3.27, 1.488, 1.185, 1.083, 1.036, and 1.018. With these values, the n-bit nonuniform quantizer produced approximately the same resolution for very small differences as an (n+2)-bit uniform quantizer, for n above 2.

3.3.2 Experiment Design, Participants, and Materials. The experiment design was similar to that in Section 3.2, but the tracking controlled compression level and this experiment involved 40 total trials: 20 with uniform quantization and 20 with nonuniform quantization, interleaved to prevent order effects.

Thirteen male subjects participated in the experiment, with age ranging from 21 to 43 years and a median age of 25. Subjects were not compensated.

Materials included the RM glove and a standard PC that was used to provide instructions to subjects, gather responses, and perform force rendering.

3.3.3 Procedure. Before the experiment, each subject practiced sample trials to help the subject understand the procedure and practice identifying compression. A minimum of three trials was required, and the subject was required to correctly detect 2-bit compression.

The subject then completed the 40 experiment trials. In the first trial of each of the two interleaved sets, 2-bit compression was used. Each correct identification resulted in compression being changed to use one more bit, up to a maximum of 8 bits. Each incorrect identification resulted in an opposite change, with a minimum of 2 bits. Software indicated to subjects whether or not responses were correct.

Table 2. Quantization experiment results

bits	2	3	4	5	6	7	8
Uniform quantizer results:							
total observations	55	57	54	44	21	13	16
correct responses	40	29	35	17	10	6	8
Nonuniform quantizer results:							
total observations	63	56	49	47	30	12	3
correct responses	39	29	30	24	12	2	1

3.3.4 Results and Discussion. Data were pooled and are shown in Table 2. Recall the natural rate of correct responses is 50%. Logit regressions were conducted as shown in Figure 3. Factor values were transformed using 2^{8-n} before the regression, and results were transformed back for plotting (n is the number of bits, and 2^{8-n} is proportional to the quantization interval in an n-bit uniform encoding).

Results suggest compression was difficult for users to detect except at the lowest resolutions. Some subjects commented on sensing possible compression artifacts in what turned out to be an uncompressed stimulus. Perceivable control changes, friction, or other noise might be interpreted as compression artifacts. Figure 4 illustrates pressure control that results for slowly squeezing a virtual spring (without compression). Spikes at the beginning of control changes should be ignored – the pressure sensor was located at a control valve where pressure changes can appear exaggerated while a valve is open. Pressure changes look roughly like a staircase in certain regions, which could produce effects indistinguishable from some compression artifacts. One of the main artifacts that may be expected with heavy compression is noticeable discrete steps due to a loss of resolution.

Results suggest the earlier choice of a 5-bit nonuniform encoding in [2] was reasonable for reducing required bandwidth without significant compression artifacts. It can be expected that users will rarely be distracted by compression artifacts with this encoding. In [2], the compression resulted in a significant increase in update rate and a reduction in latency. When haptic feedback quality is limited by communication bandwidth, the net effect of a properly chosen coding will be improvement in overall quality due to improved update rate and latency.

Nonuniform quantization appeared slightly better than uniform quantization, but further investigation is needed for any strong conclusions and to investigate the effects of changing parameter value *b*.

In conclusion, device characteristics allowed heavy compression with minimal detectable artifacts. Only small changes needed to be represented due to small per-cycle differences, and these already small changes did not need to be represented at a high resolution.

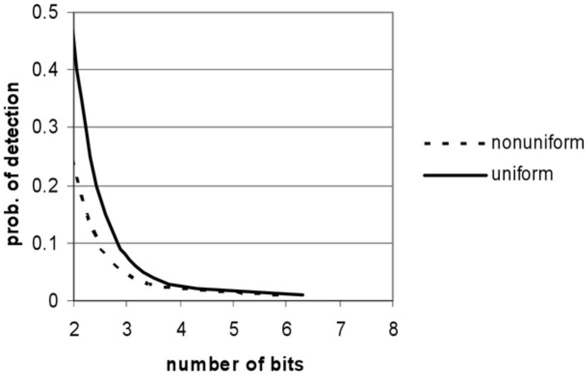

Fig. 3. Logit regressions (Pearson goodness-of-fit: uniform - $X^2(5)$ = 6.9, P=0.23, nonuniform - $X^2(5)$ = 9.22, P=0.10)

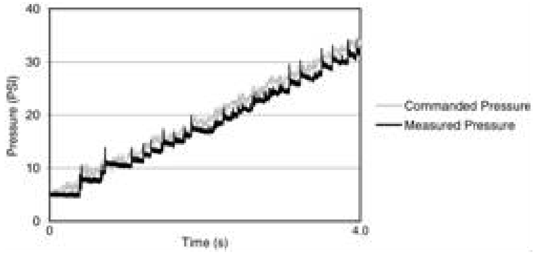

Fig. 4. Pressure measured at control valve

4. Conclusion and Future Work

A predictive coding was presented and evaluated for compression in a force-feedback display. The evaluation considered encoding of a sensor reading, but similar results would be expected with encoding of force levels, since, in this case, change in force was simply proportional to change in the sensor value.

The paper focused mainly on a simple first-order difference. The potential for further compression with a higher-order function was shown in Section 3.1 but not considered in detail. Development of good higher-order functions for haptic data is one of several issues that should be addressed in the future. Other areas of study include adaptive sampling or encoding techniques, identification of the best nonuniform quantizers, and development of a stronger understanding of the tradeoffs between various system parameters (update rate, compression rate, noise filter characteristics, latency, etc.) so these can be optimized to provide the best user experience. Results should be related to both psychophysical responses and haptic device characteristics such as mechanical bandwidth, so ultimately an overall model can emerge to guide the design of effective future haptic interfaces. Finally, differences between users should be considered by future evaluations. Some users will be more sensitive to compression artifacts than others. A compression system designed to be effective for an average user may be unpleasant for experienced users.

Acknowledgments

The author thanks Matthew G. Price for his work on experiment software and administering experiments during an undergraduate project. The author also thanks the experiment subjects.

References

[1] C. W. Borst, "Haptic Feedback for Virtual or Mixed Reality Environments using a Combination of Passive Haptics and Active Force Feedback," PhD dissertation, Dept. of Computer Science. Texas A&M University, College Station, 2002.

[2] C. W. Borst and R. A. Volz, "Observations on and Modifications to the Rutgers Master to Support a Mixture of Passive Haptics and Active Force Feedback," Eleventh Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Los Angeles, 2003.

[3] A. Ortega and Y. Liu, "Lossy Compression of Haptic Data," in *Touch in Virtual Environments: Haptics and the Design of Interactive Systems*, M. L. McLaughlin, J. P. Hespanha, and G. S. Sukhatme, Eds. Upper Saddle River, New Jersey: Prentice Hall PTR, 2002.

[4] C. Shahabi, A. Ortega, and M. R. Kolahdouzan, "A comparison of different haptic compression techniques," IEEE International Conference on Multimedia and Expo, 2002.

[5] M. Bouzit, G. Burdea, G. Popescu, and R. Boian, "The Rutgers Master II-New Design Force-Feedback Glove," *IEEE/ASME Transactions on Mechatronics*, vol. 7, pp. 256-263, 2002.

Appendix

The following figures are absent from the printed proceedings but are included in the Conference CD-ROM proceedings.

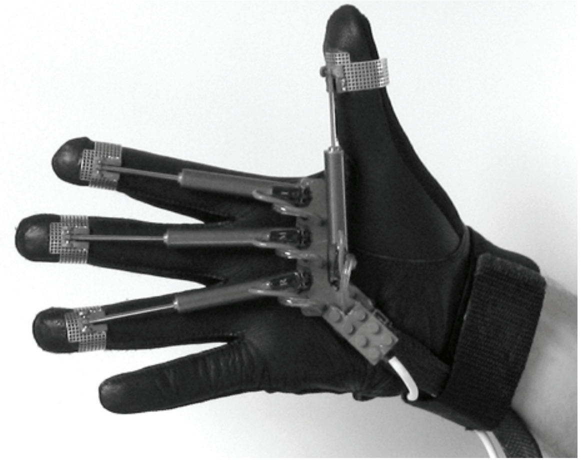

Fig. A1. Pneumatic force-feedback glove used in the experiments

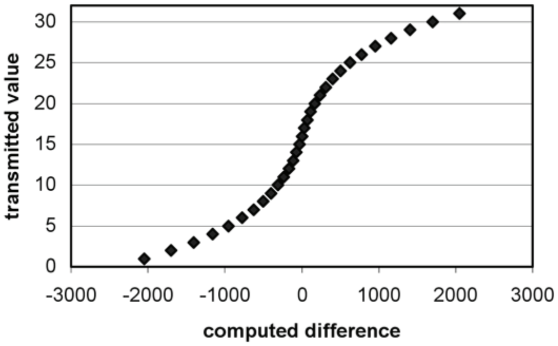

Fig. A2. Example of a nonuniform coding, showing the 5-bit quantizer/encoder